



TUGAS AKHIR - SS141501

PERBANDINGAN KINERJA METODE *FUZZY K-MEANS CLUSTERING* DAN *FUZZY GUSTAFSON-KESSEL CLUSTERING* BERDASARKAN REALISASI PAJAK DAERAH KOTA SURABAYA

**YANUAR RAFI RAHADIAN
NRP 1313 100 091**

**Dosen Pembimbing
Raden Mohamad Atok, Ph.D.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



TUGAS AKHIR - SS141501

PERBANDINGAN KINERJA METODE *FUZZY K-MEANS CLUSTERING* DAN *FUZZY GUSTAFSON-KESSEL CLUSTERING* BERDASARKAN REALISASI PAJAK DAERAH KOTA SURABAYA

**YANUAR RAFI RAHADIAN
NRP 1313 100 091**

**Dosen Pembimbing
Raden Mohamad Atok, Ph.D.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



FINAL PROJECT - SS141501

**PERFORMANCE COMPARISON OF FUZZY K-MEANS
CLUSTERING AND FUZZY GUSTAFSON-KESSEL
CLUSTERING BASED ON REGIONAL TAX REALIZATION
IN SURABAYA CITY**

**YANUAR RAFI RAHADIAN
NRP 1313 100 091**

**Supervisor
Raden Mohamad Atok, Ph.D.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

LEMBAR PENGESAHAN

PERBANDINGAN KINERJA METODE FUZZY K- MEANS CLUSTERING DAN FUZZY GUSTAFSON- KESSEL CLUSTERING BERDASARKAN REALISASI PAJAK DAERAH KOTA SURABAYA

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

Oleh :

Yanuar Rafi Rahadian

NRP. 1313 100 091

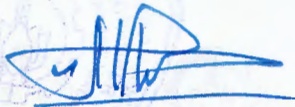
Disetujui oleh Pembimbing:

Raden Mohamad Atok, Ph.D.

NIP. 19710915 199702 1 001

()

Mengetahui,
Kepala Departemen



Dr. Suhartono

NIP. 19710929 199512 1 001



SURABAYA, 15 JANUARI 2018

PERBANDINGAN KINERJA METODE *FUZZY K-MEANS CLUSTERING* DAN *FUZZY GUSTAFSON-KESSEL CLUSTERING* BERDASARKAN REALISASI PAJAK DAERAH KOTA SURABAYA

Nama Mahasiswa : Yanuar Rafi Rahadian
NRP : 1313 100 091
Departemen : Statistika
Dosen Pembimbing : Raden Mohamad Atok, Ph.D.

Abstrak

Kota Surabaya merupakan kabupaten/kota dengan pendapatan asli daerah tertinggi di Indonesia. Di sisi lain, Surabaya memiliki masalah ketimpangan wilayah yang tinggi berdasarkan volume PDRB per kecamatan. Penelitian ini bertujuan mengelompokkan realisasi pajak daerah tingkat kecamatan di Kota Surabaya. Studi simulasi pengelompokan menunjukkan bahwa akurasi metode Fuzzy Gustafson-Kessel (FGK) cenderung lebih tinggi dibandingkan metode Fuzzy K-Means (FKM). Digunakan 7 variabel pajak daerah di Kota Surabaya, yaitu pajak air tanah, pajak hiburan, pajak penerangan jalan, pajak hotel, pajak restoran, pajak parkir serta pajak bumi dan bangunan. Pereduksian variabel menggunakan analisis faktor menghasilkan 2 komponen utama yang menjelaskan varians total sebesar 73%. Jumlah optimum klaster yang dihasilkan metode FGK adalah 4 sedangkan metode FKM adalah 2 klaster. Perbandingan antara hasil optimum metode FGK dan metode FKM menghasilkan metode FGK yang lebih akurat. Jumlah pajak kecamatan tertinggi ditunjukkan oleh rata-rata tertinggi pada klaster yaitu kecamatan Dukuh Pakis, Gayungan, Genteng, Mulyorejo, Sambikerep, Tegalsari dan Wonokromo.

Kata kunci: Realisasi pajak daerah, analisis faktor, fuzzy k-means, fuzzy Gustafson-Kessel.

(Halaman ini sengaja dikosongkan)

PERFORMANCE COMPARISON OF FUZZY K-MEANS CLUSTERING AND FUZZY GUSTAFSON-KESSEL CLUSTERING BASED ON REGIONAL TAX REALIZATION IN SURABAYA CITY

Student Name : Yanuar Rafi Rahadian
Student Number : 1313 100 091
Department : Statistics
Supervisor 1 : Raden Mohamad Atok, Ph.D.

Abstract

Surabaya city is the highest regency income in Indonesia. On the other hand, Surabaya city has a high regional inequality problem based on GDP total between sub-districts. The objective of this research is to classify the realization of regional taxes in each sub-district in Surabaya. Simulation study shows that Fuzzy Gustafson-Kessel (FGK) method more accurate than Fuzzy K-Means (FKM) method. This research uses 7 variables of regional tax in Surabaya. Factor analysis is used to reduce 7 regional tax variables into 2 main components that can explain the total variance of 73%. FGK method reach 4 clusters as optimum number of clusters, meanwhile 2 clusters for FKM method. Comparing the two results, FGK method is more accurate than FKM method based on icd-rate criteria. The highest tax rate indicated by sub-districts of the highest average cluster that is Dukuh Pakis, Gayungan, Genteng, Mulyorejo, Sambikerep, Tegalsari and Wonokromo.

Keywords: *Realization of regional tax, factor analysis, fuzzy k-means, fuzzy Gustafson-Kessel.*

KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah memberi kenikmatan, rizki, kemudahan serta karunia-Nya. Shalawat serta salam senantiasa tercurahkan kepada junjungan Nabi Muhammad SAW. Semoga Allah SWT senantiasa memberi perlindungan, kesehatan, dan waktu sehingga dapat selalu mendekatkan diri kepadaNya.

Terselesaikannya Tugas Akhir ini dengan judul **“PERBANDINGAN KINERJA METODE *FUZZY K-MEANS CLUSTERING* DAN *FUZZY GUSTAFSON-KESSEL CLUSTERING* BERDASARKAN REALISASI PAJAK DAERAH KOTA SURABAYA”** dengan baik tak lepas peran bantuan, arahan, dan petunjuk berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih dengan penuh hormat dan kerendahan hati kepada:

1. Raden Mohamad Atok, Ph.D. selaku dosen pembimbing yang telah sabar membimbing dan mengarahkan sepenuh hati kepada penulis selama penyusunan Tugas Akhir.
2. Dr. Suhartono dan Dr. rer. pol. Dedy Dwi Prastyo selaku dosen penguji yang telah banyak memberikan masukan dalam penyusunan Tugas Akhir.
3. Kedua orang tua, Rudy Setiawan dan Zamrud Mufida yang selalu memberikan dukungan, doa, materi, kesabaran, dan pengertian yang tiada batas yang telah diberikan serta adik Johan Nabiel Raihan yang telah memberikan keceriaan.
4. Sahabat yang selalu memberi motivasi, dorongan, menemani, dan bertukar pikiran dalam penyusunan Tugas Akhir ini, Aya, Adit, Azhar, Vina, Jefri, Bimbim, Bani, dll.
5. Teman-teman mahasiswa S1 Statistika ITS angkatan 2013 yang tidak dapat penulis sebutkan satu per satu.
6. Roliyono, S.E., M.M. selaku pembimbing lapangan dan seluruh karyawan Badan Pengelolaan Keuangan dan Pajak Daerah Kota Surabaya yang telah bersedia membantu dalam proses pengambilan data.

7. Seluruh tenaga kependidikan Departemen Statistika FMKSD ITS yang membantu keperluan bidang akademik.
8. Seluruh civitas akademika ITS dan pihak-pihak yang telah membantu penyusunan Tugas Akhir, baik secara langsung maupun tidak langsung.

Dengan berakhirnya pengerjaan laporan ini, penulis berharap agar laporan ini dapat memberikan manfaat kepada penulis, pembaca dan instansi terkait. Penulis menyadari bahwa laporan ini masih jauh dari kesempurnaan, oleh karena itu kritik dan saran yang membangun diharapkan untuk pengembangan Tugas Akhir ini.

Surabaya, 31 Desember 2017

Penulis

(Halaman ini sengaja dikosongkan)

DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN.....	v
ABSTRAK.....	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI.....	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL.....	xvii
DAFTAR LAMPIRAN	xix

BAB I PENDAHULUAN

1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan.....	6
1.4 Manfaat.....	6
1.5 Batasan Masalah.....	6

BAB II TINJAUAN PUSTAKA

2.1 Statistika Deskriptif	7
2.2 Pengujian Distribusi Normal Multivariat	7
2.3 Pengujian Kecukupan Data	8
2.4 Pengujian Independensi.....	9
2.5 Analisis Faktor	9
2.5.1 Analisis Komponen Utama.....	10
2.5.2 Penentuan Jumlah Faktor.....	11
2.5.3 Rotasi Faktor.....	11
2.6 Analisis Klaster	11
2.7 Fungsi Keanggotaan	12
2.8 Analisis Klaster FKM.....	14
2.9 Analisis Klaster FGK	16
2.10 <i>Calinski Harabasz Pseudo F-Statistic</i>	17
2.11 Evaluasi Kinerja Pengelompokan	18
2.12 Profil Kota Surabaya	18
2.13 Pengertian Pajak.....	20
2.14 Pengertian Pajak Daerah	21

2.15 Studi Simulasi Pengelompokan.....	22
2.16 Penelitian Sebelumnya	23
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	25
3.2 Variabel Penelitian	25
3.3 Definisi Operasional.....	26
3.4 Desain Studi Simulasi Pengelompokan.....	28
3.5 Langkah Analisis	31
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Studi Simulasi Pengelompokan Kecamatan.....	35
4.2 Karakteristik Pajak Daerah di Kota Surabaya	37
4.3 Pra-Pengelompokan Kecamatan di Kota Surabaya...	38
4.4 Pengelompokan Kecamatan di Kota Surabaya.....	41
4.5 Evaluasi Kinerja Metode Pengelompokan	45
4.6 Karakteristik Klaster di Kota Surabaya.....	46
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan.....	51
5.2 Saran.....	52
DAFTAR PUSTAKA	53
LAMPIRAN.....	57

DAFTAR GAMBAR

Gambar 2.1	Representasi Fungsi Linier Naik	13
Gambar 2.2	Peta Lahan Kota Surabaya.....	19
Gambar 3.1	Diagram Alir Penelitian.....	33
Gambar 4.1	Akurasi FGK-FKM Normal Multivariat $K=2$...	33
Gambar 4.2	Akurasi FGK-FKM Normal Univariat $K=2$	34
Gambar 4.3	Peta Pengelompokan Kecamatan di Surabaya...	47
Gambar 4.4	<i>Boxplot</i> Komponen Faktor 1 Kota Surabaya	48
Gambar 4.5	<i>Boxplot</i> Komponen Faktor 2 Kota Surabaya	48

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

Tabel 3.1	Struktur Data Penelitian.....	25
Tabel 3.2	Unit Penelitian	25
Tabel 3.3	Variabel Penelitian.....	26
Tabel 4.1	Perbandingan Akurasi Normal Multivariat.....	36
Tabel 4.2	Karakteristik Pajak Daerah Surabaya	37
Tabel 4.3	<i>Communalites</i> Kecamatan Surabaya.....	40
Tabel 4.4	Proporsi Keragaman Kecamatan Surabaya.....	41
Tabel 4.5	Matriks Komponen	42
Tabel 4.6	Matriks Komponen Setelah Rotasi	42
Tabel 4.7	Hasil Pengelompokan FKM Kota Surabaya	43
Tabel 4.8	Hasil Pengelompokan FGK Kota Surabaya.....	44
Tabel 4.9	Perhitungan <i>Pseudo F</i> Kota Surabaya	45
Tabel 4.10	Profil Pengelompokan Kecamatan.....	45
Tabel 4.11	Perbandingan Karakteristik Klaster Surabaya	49

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Lampiran 1.	Data Penelitian	57
Lampiran 2.	<i>Output KMO dan Bartlett Test</i>	64
Lampiran 3.	<i>Output Nilai Communalities</i>	64
Lampiran 4.	<i>Output Scree Plot</i>	64
Lampiran 5.	Tabel <i>Membership Function</i>	65
Lampiran 6.	Simulasi FGK Normal Multivariat.....	71
Lampiran 7.	Simulasi FGK Normal Univariat.....	74
Lampiran 8.	Simulasi FKM Normal Multivariat	77
Lampiran 9.	Simulasi FKM Normal Univariat	80
Lampiran 10.	Tabel Akurasi Normal Multivariat	83
Lampiran 11.	Tabel Akurasi Normal Multivariat <i>c</i> Besar.	83
Lampiran 12.	Tabel Akurasi Normal Univariat.....	84
Lampiran 13.	Tabel Akurasi Normal Univariat <i>c</i> Besar.	84
Lampiran 14.	Hasil Pengelompokan FKM Surabaya.	85
Lampiran 15.	Hasil Pengelompokan FGK Surabaya.....	86
Lampiran 16.	Plot Pengelompokan FKM Surabaya.	87
Lampiran 17.	Plot Pengelompokan FGK Surabaya.....	88
Lampiran 18.	<i>Syntax Uji Multivariat Normal</i>	89
Lampiran 19.	<i>Syntax Pseudo F-Statistics & Icd-rate</i>	89
Lampiran 20.	<i>Syntax FKM dan FGK Normal Multivariat</i>	91
Lampiran 21.	<i>Syntax FKM dan FGK Normal Univariat</i>	91
Lampiran 22.	<i>Syntax FKM dan FGK Surabaya</i>	92
Lampiran 23.	Surat Pernyataan Data	93

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Terwujudnya masyarakat adil dan makmur melalui berbagai upaya peningkatan taraf hidup kesejahteraan rakyat merupakan salah satu cita-cita pembangunan nasional Republik Indonesia (Kusuma, 1992). Pemerintah dan masyarakat telah diatur melalui perundangan untuk menyelenggarakan pembangunan nasional secara bersama-sama. Anggaran dana merupakan salah satu instrumen penting dalam berjalannya roda pemerintahan dan pembangunan nasional. Anggaran Pendapatan dan Belanja Negara (APBN) setiap tahunnya disusun sebagai landasan anggaran resmi yang diumumkan negara dengan dua unsur utama yaitu penerimaan (*revenue*) dan pengeluaran (*expenditure*). Kebutuhan dana pembangunan nasional setiap tahunnya meningkat seiring dengan peningkatan jumlah dan kebutuhan masyarakat sehingga dana yang dibutuhkan untuk pembangunan negara sangat besar (BPPK, 2015).

Peningkatan kebutuhan dana pembangunan nasional harus selalu dicapai melalui peningkatan pendapatan asli daerah. Undang-Undang Republik Indonesia Nomor 28 Tahun 2007 pasal 1 ayat 1 menyatakan bahwa pajak merupakan kontribusi wajib kepada negara yang terutang oleh orang pribadi atau badan yang bersifat memaksa berdasarkan undang-undang tanpa mendapatkan imbalan langsung. Pajak merupakan salah satu instrumen pendapatan asli daerah (BPPK, 2015). Pembiayaan kepentingan umum sebagian besar diambil dari pajak yang dipungut oleh pemerintah untuk diberikan kepada masyarakat dalam mencapai tujuan pembangunan nasional (Nurcholish, 2005). Pemerintah pusat memperluas wewenang keuangan kepada pemerintah daerah melalui otonomi daerah yang diatur oleh perundangan yang berlaku sehingga pemerintah daerah memiliki kewenangan memungut pajak daerah (Dirjen Anggaran Kemenkeu, 2011).

Kota Surabaya merupakan pusat pemerintahan Provinsi Jawa Timur dengan tingkat pendapatan asli daerah tertinggi di Indonesia (Dirjen Perimbangan Kemenkeu, 2017). Realitas yang mengikat pada kota Surabaya ini memiliki konsekuensi yang besar terhadap laju pemerataan ekonomi wilayah yang selaras dengan upaya peningkatan kemampuan pemerintah daerah dalam mengelola pajak daerah. Manfaat lain yang berpotensi adalah peningkatan aksesibilitas yang didukung oleh saran dan prasarana, transportasi dan komunikasi (Adika, 2008). Berdasarkan peraturan daerah Kota Surabaya, pajak daerah terdiri dari pajak air tanah, pajak reklame, pajak penerangan jalan, pajak hotel, pajak restoran, pajak hiburan, pajak parkir, serta pajak bumi dan bangunan. Kota Surabaya memiliki 31 kecamatan yang memiliki karakteristik dan potensi pendapatan pajak yang berbeda-beda. Kondisi tipologi setiap kecamatan yang berbeda-beda ini perlu dimonitoring dengan melakukan pengumpulan data terkait karakteristik pajak daerah.

Ilmu pengetahuan yang berhubungan dengan metode mengumpulkan data, menganalisis data, menginterpretasi hasil maupun ilmu yang dapat mempresentasikan kesimpulan analisis adalah ilmu statistika. Terdapat variabel penelitian dalam data penelitian yang dapat diidentifikasi tingkat struktur hubungan antar variabel dengan analisis faktor. Setelah struktur hubungan antar variabel telah dianalisis maka data akan semakin mudah dikelompokkan (Simamora, 2005). Pengelompokan suatu gugus data penelitian bertujuan untuk menentukan suatu kelompok yang alami dari kelompok-kelompok individu atau membagi data menjadi beberapa kelompok bersifat *unsupervised learning* (Rencher, 2002). Kelompok individu ini bisa membentuk populasi yang lengkap atau suatu sampel dari populasi yang lebih besar. Salah satu metode pengelompokan yang populer adalah metode *k-means clustering*. Algoritma dari metode klasik ini akan membagi data atau obyek ke dalam k buah kelompok dan pada setiap *cluster* yang terbentuk akan selalu terdapat *centroid* yang

merepresentasikan anggota maupun sifat dari suatu *cluster* tersebut.

Zadeh (1965) mengembangkan *k-means clustering* berlandaskan logika *fuzzy*. Logika *fuzzy* ini berlandaskan pada ide bahwa A dapat sama dengan bukan A. Zadeh (1965) mengamati bahwa teori himpunan konvensional tidak mencukupi untuk mengatasi proses perubahan yang halus. Perubahan yang halus lebih bersifat perubahan yang berangsur-angsur sehingga penjelasan himpunan kelompok harus dideskripsikan dalam beberapa tahapan. Logika *fuzzy* menyediakan teori kebenaran *fuzzy* yang dinyatakan dalam tingkat derajat keanggotaan A atau bukan A dalam suatu semesta permasalahan.

Fuzzy clustering yaitu suatu metode yang melakukan pengelompokan dengan mempertimbangkan tingkat derajat keanggotaan mencakup himpunan *Fuzzy* sebagai landasan pembobotan pengelompokan. *Fuzzy k-means* (FKM) *clustering* dan *fuzzy Gustafson-Kessel* (FGK) merupakan contoh pengelompokan dengan *fuzzy*.

Analisis pengelompokan yang digunakan dalam metode FKM ini menggunakan jarak *euclidean* sebagai *default* ukuran jaraknya, dimana ukuran jarak ini hanya bisa mendeteksi kelompok dengan susunan *spherical*. Berdasarkan referensi peneliti utama oleh Gustafson dkk. (1978), FGK *clustering* merupakan pengembangan dari metode FKM *clustering* yang tidak hanya membuat setiap titik obyek masuk ke lebih dari satu klaster, namun metode ini juga mampu mengelompokan data berdasarkan struktur klaster yang berbentuk *spherical*, *elipsoid* atau bentuk yang lain. Pengelompokan berdasarkan FGK *clustering* menawarkan bentuk modifikasi pada komponen jarak dalam fungsi tujuan yang diminimasi menurut Hartati dkk. (2005).

Terkait keterhubungan pajak dengan klaster, Velichkov dkk. (2017) telah menjelaskan bahwa sistem pajak memiliki efek terhadap kestabilan makroekonomi dimana analisis klaster merupakan metode pilihan untuk mengetahui gambaran keadaan dan

pengelompokan pajak. Oleh karena itu, peneliti menerapkan analisis kluster dalam memonitoring pajak dengan menggunakan metode FGK dan FKM pada kecamatan di Kota Surabaya. Kemudian, kinerja kedua metode dalam melakukan pengelompokan dapat dievaluasi melalui *icd-rate* menurut Mingoti dkk. (2016).

Penelitian terdahulu oleh Sirait (2015) menyimpulkan bahwa PDRB, jumlah hotel, jumlah pelanggan PLN, jumlah wisatawan dan kepadatan penduduk berpengaruh signifikan terhadap pajak daerah kota Surabaya. Selain itu, terdapat dua kelompok pajak daerah yang terbentuk dari penggunaan analisis kluster hierarki dengan metode *ward* (Lutfi, 2016) dan penggunaan metode *fuzzy c-shell* dan metode FKM berdasarkan variabel pembentuk IPM (Widodo, 2012). Riset terkait metode yang digunakan pada penelitian ini telah dilakukan oleh Amalia dkk. (2016) tentang Analisis Clustering Perusahaan Sub Sektor Perbankan berdasarkan Rasio Keuangan CAMELS Tahun 2014 menggunakan metode FKM dan FGK dengan hasil diperoleh kondisi optimum untuk kedua metode berdasarkan nilai Index XB sebanyak 2 *cluster* dengan metode terbaik adalah FGK berdasarkan nilai *icd-rate* terkecil yaitu sebesar 0,7232. Berdasarkan jumlah *cluster* optimum dan metode terbaik, hasil pengujian MANOVA menghasilkan bahwa tidak ada perbedaan karakteristik antar cluster terhadap respon menurut Amalia dkk. (2016).

Berdasarkan uraian sebelumnya, maka pada penelitian ini terdapat tiga topik yang akan dibahas. Topik pertama adalah analisis deskriptif untuk mendeskripsikan karakteristik realisasi pajak daerah di Kota Surabaya. Studi simulasi pengelompokan menjadi topik kedua penelitian. Topik ketiga diawali dengan analisis faktor untuk mereduksi variabel pajak daerah menjadi faktor komponen. Kemudian analisis kluster FGK dan FKM dilakukan dengan penentuan kluster optimum dan kluster terbaik. Karakteristik tiap kluster dapat dilihat melalui *boxplot*. Dengan demikian penelitian ini diharapkan mampu memberikan informasi

statistik kepada Badan Pengelolaan Keuangan dan Pajak Daerah Kota Surabaya.

1.2 Rumusan Masalah

Permasalahan yang dirumuskan dalam penelitian ini adalah sebagai berikut:

1. Bagaimana akurasi pengelompokan metode FKM *clustering* dan FGK *clustering*?
2. Bagaimana karakteristik realisasi pajak daerah tiap kecamatan di Kota Surabaya Tahun 2016?
3. Bagaimana pengelompokan realisasi pajak daerah tiap kecamatan berdasarkan metode FKM *clustering* dan FGK *clustering*?
4. Bagaimana evaluasi kinerja metode FKM *clustering* dan FGK *clustering* terhadap setiap kluster realisasi pajak daerah yang terbentuk?
5. Bagaimana karakteristik kluster yang terbentuk di Kota Surabaya Tahun 2016?

1.3 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengetahui akurasi pengelompokan metode FKM dan FGK dalam mengelompokkan data berdistribusi tertentu.
2. Mendeskripsikan karakteristik realisasi pajak daerah tiap kecamatan di Kota Surabaya Tahun 2016.
3. Mengelompokkan realisasi pajak daerah tiap kecamatan berdasarkan metode FKM *clustering* dan FGK *clustering*.
4. Melakukan analisis evaluasi kinerja metode FKM *clustering* dan FGK *clustering* terhadap setiap kluster realisasi pajak daerah yang terbentuk.
5. Mendeskripsikan karakteristik kluster yang terbentuk di Kota Surabaya Tahun 2016.

1.4 Manfaat

Manfaat penelitian ini diharapkan dapat memberikan informasi mengenai akurasi metode, karakteristik realisasi pajak

daerah per kecamatan, hasil dan evaluasi pengelompokan realisasi pajak daerah kecamatan di Kota Surabaya kepada Badan Pengelolaan Keuangan dan Pajak Daerah (BPKPD) Kota Surabaya dan penelitian selanjutnya.

1.5 Batasan Masalah

Pada penelitian ini terdapat batasan masalah yaitu hanya dua komponen berdasarkan teori nilai kumulatif proporsi keragaman yang digunakan dalam analisis faktor. Analisis faktor digunakan sebagai salah satu pendekatan dalam mereduksi variabel awal. Batasan masalah kedua adalah inisiasi klaster yang dibandingkan untuk memperoleh klaster optimum sebanyak 2 klaster, 3 klaster dan 4 klaster. Pemilihan tiga inisiasi jumlah klaster ditentukan untuk memudahkan pengambilan kesimpulan akhir. Kusuma (2017) menyimpulkan bahwa klaster yang dihasilkan dari fungsi keanggotaan linier naik optimum pada 6 provinsi dari 9 provinsi di Sumatera. Sehingga Batasan masalah ketiga terletak pada penggunaan fungsi keanggotaan linier naik.

BAB II TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistika deskriptif adalah metode statistika yang berfungsi untuk memberikan gambaran umum tentang penyajian sampel dan populasi. Analisis statistika deskriptif dapat diartikan sebagai metode yang berkaitan dengan mengumpulkan, mengolah, dan menyajikan data sehingga memberikan informasi yang berguna (Walpole, 1998). Data dapat disajikan dalam bentuk grafik atau tabel, untuk ukuran pemusatan data dan ukuran penyebaran data dideskripsikan secara numerik. Ukuran pemusatan data meliputi rata-rata, nilai tengah, dan modus sedangkan ukuran penyebaran data meliputi ragam dan standar deviasi menurut Walpole dkk. (2012).

2.2 Pengujian Distribusi Normal Multivariat

Pada penerapan analisis multivariat, asumsi normal multivariat diperlukan untuk memastikan keadaan input data mengikuti distribusi normal multivariate atau tidak. Pemeriksaan asumsi normal multivariat dapat dilakukan dengan pengujian koefisien korelasi antara nilai mahalanobis dengan nilai *quantile chi-square*. Untuk mengetahui hal tersebut dapat dilakukan pengujian melalui nilai koefisien korelasi dengan hipotesis sebagai berikut (Johnson dan Wichern, 2007).

H_0 : Data berdistribusi normal multivariat

H_1 : Data tidak berdistribusi normal multivariat

Dengan rumus statistik uji yang digunakan seperti yang ditunjukkan pada Persamaan (2.1).

$$r_Q = \frac{\sum_{i=1}^c (d_i - \bar{d})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^c (d_i - \bar{d})^2} \sqrt{\sum_{i=1}^c (q_i - \bar{q})^2}} \quad (2.1)$$

dimana:

r_Q = koefisien korelasi antara jarak mahalanobis yang telah diurutkan dengan nilai *quantile Chi-Square*

q_i = *quantile Chi-Square*

p = banyaknya variabel

c = banyaknya pengamatan

$i = 1, 2, 3, \dots, c$

d_i = jarak mahalanobis

\bar{d} = nilai tengah jarak mahalanobis

\bar{q} = nilai tengah *quantile Chi-Square*

Berdasarkan Hair dkk. (2009), apabila didapatkan nilai statistik uji r_Q lebih besar dari nilai r_{tabel} maka dapat disimpulkan gagal tolak H_0 , yang berarti bahwa data memenuhi asumsi distribusi normal multivariat.

2.3 Pengujian Kecukupan Data

Measurement of Sampling Adequacy (MSA) adalah indeks perbandingan jarak antara koefisien korelasi dengan koefisien korelasi parsialnya. Berdasarkan Hair dkk. (2009), apabila kuadrat dari elemen matriks korelasi **D** di antara seluruh pasangan variabel bernilai lebih kecil jika dibandingkan dengan jumlah kuadrat dari elemen matriks korelasi **R**, maka akan menghasilkan nilai MSA mendekati 1.

Nilai MSA dianggap mencukupi jika lebih dari 0,5. Analisis MSA berfungsi untuk menguji kecukupan data yang telah terambil meliputi segi kecukupan data untuk difaktorkan. Rumus yang digunakan ditunjukkan pada Persamaan (2.2).

$$MSA = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p q_{ij}^2} \quad (2.2)$$

dimana:

r_{ij}^2 = kuadrat dari elemen matriks korelasi \mathbf{R}

q_{ij}^2 = kuadrat dari elemen matriks korelasi $\mathbf{D} = \mathbf{R}^{-1}\mathbf{D}$

$i, j = 1, 2, 3, \dots, p$

2.4 Pengujian Independensi

Uji independensi digunakan untuk mengetahui besarnya nilai korelasi antar variabel. Apabila variabel x_1, x_2, \dots, x_N independen (bersifat saling bebas), maka matriks korelasi antar variabel sama dengan matriks identitas (Rencher, 1998). Berikut ini adalah hipotesis yang digunakan dalam *Bartlett Test*.

$H_0 : \boldsymbol{\rho} = \mathbf{I}$ (tidak terdapat korelasi multivariat)

$H_1 : \boldsymbol{\rho} \neq \mathbf{I}$ (terdapat korelasi multivariat)

Dengan statistik uji ditunjukkan pada Persamaan (2.3).

$$\chi^2 = - \left(c - 1 - \frac{2N+5}{6} \right) \ln|\mathbf{R}| \quad (2.3)$$

dimana:

$|\mathbf{R}|$ = nilai determinan dari matriks korelasi

N = banyak variabel

Adapun H_0 akan ditolak jika $\chi^2 > \chi_{\alpha, \frac{1}{2}N(N-1)}^2$ yang berarti terdapat korelasi multivariat antar variabel penelitian (Rencher, 1998).

2.5 Analisis Faktor

Analisis Faktor adalah suatu analisis data untuk mengetahui faktor-faktor yang dominan dalam menjelaskan suatu masalah. Misalkan x adalah variabel random dengan c komponen memiliki vektor mean $\boldsymbol{\mu}$ dan matriks kovarians $\boldsymbol{\Sigma}$ dimana masing-masing variabel pada model analisis faktor merupakan kombinasi linier dari skor faktor f_1, f_2, \dots, f_N dengan disertai error untuk menghitung bagian dari variabel yang unik atau tidak sama dengan variabel yang lain. Adapun model untuk x_1, x_2, \dots, x_c pada persamaan vektor \mathbf{x} ditunjukkan pada Persamaan (2.4) (Rencher, 1998).

$$\begin{aligned}
x_1 - \mu_1 &= \lambda_{11}f_1 + \lambda_{12}f_2 + \cdots + \lambda_{1N}f_N + \varepsilon_1 \\
x_2 - \mu_2 &= \lambda_{21}f_1 + \lambda_{22}f_2 + \cdots + \lambda_{2N}f_N + \varepsilon_2 \\
&\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\
x_c - \mu_c &= \lambda_{c1}f_1 + \lambda_{c2}f_2 + \cdots + \lambda_{cN}f_N + \varepsilon_c
\end{aligned} \tag{2.4}$$

dimana :

μ_i = rata-rata variabel ke- λ ; $i = 1, 2, \dots, c$

λ_{ij} = *loading* pengamatan ke- i pada faktor ke- j ; $i = 1, 2, \dots, c$

ε_i = faktor spesifik ke- λ ; $i = 1, 2, \dots, c$

f_j = komponen faktor ke- j ; $j = 1, 2, \dots, N$

Idealnya N harus jauh lebih kecil daripada c . Koefisien λ_{ij} disebut *loading* dan berfungsi sebagai bobot yang menunjukkan bagaimana setiap f_j secara individual dipengaruhi oleh komponen faktor f_1, f_2, \dots, f_N . Dengan kata lain, λ_{ij} menunjukkan pentingnya faktor f_j terhadap variabel x_i dan bisa digunakan untuk menginterpretasikan faktor f_j . Setelah mengestimasi λ_{ij} dan merotasinya, diharapkan variabel-variabel bisa terpartisi ke dalam pengelompokan sesuai faktor yang terbentuk. Variabel dengan bobot *loading* terbesar pada tiap komponen faktor yang terbentuk dapat digunakan untuk landasan analisis kluster (Sharma, 1996).

2.5.1 Analisis Komponen Utama

Analisis komponen utama merupakan salah satu pendekatan metode analisis faktor untuk mentransformasi variabel asli yang masih saling berkorelasi satu dengan yang lain menjadi satu set peubah baru yang tidak berkorelasi. Tujuan penggunaan analisis komponen utama adalah mengondisikan tiap komponen yang terbentuk dapat mewakili seluruh variabel, sehingga hasil dari komponen faktor akan diubah menjadi derajat keanggotaan. Kemudian dari derajat keanggotaan akan dibentuk matriks berukuran k objek pada pengelompokan *fuzzy k-means clustering* dan *fuzzy gustafson-kessel clustering* pada analisis selanjutnya (Hartati, 2005).

Principal component analysis menggunakan varians sebagai dasar perhitungan (Simamora, 2005). Tidak ada kesepakatan suatu ukuran besar proporsi keragaman data yang dianggap cukup

mewakili keragaman total, namun peneliti banyak menggunakan angka 70% hingga 90% (Johnson & Wichern, 2007). Nilai komponen utama atau variabel representatif dari masing-masing komponen dapat digunakan apabila analisis lanjutan dari analisis faktor adalah analisis kluster (Sharma, 1996).

2.5.2 Penentuan Jumlah Faktor

Penentuan jumlah faktor yang dibentuk dapat menggunakan tiga kriteria yang biasa digunakan dalam analisis komponen utama (Rencher, 1998). Kriteria yang paling sering digunakan pada analisis komponen utama adalah penentuan banyak faktor berdasarkan proporsi faktor ke- j dalam menjelaskan total varians sampel. Matriks korelasi yang bernilai lebih dari 1 dapat menjadi faktor berdasarkan kriteria *Eigenvalues*. Selain penentuan banyak faktor dapat dilihat secara visual melalui *scree plot*.

2.5.3 Rotasi Faktor

Setelah diketahui jumlah faktor yang terbentuk, tahapan selanjutnya adalah menginterpretasikan komponen dari faktor tersebut. Terkadang *loadings factor* yang dihasilkan memiliki nilai yang hampir sama pada setiap kolomnya, sehingga akan terjadi kesulitan dalam menginterpretasikannya. Rotasi faktor bertujuan untuk menyederhanakan struktur faktor sehingga mudah untuk diinterpretasikan. Ada dua metode rotasi yang berbeda yaitu *orthogonal* dan *oblique rotation* (Rencher, 1998). Metode rotasi yang populer ialah *varimax procedure*. Prosedur ini merupakan metode *orthogonal* yang berusaha memaksimumkan nilai kuadrat dari *loading*.

2.6 Analisis Kluster

Analisis Kluster merupakan sebuah teknik untuk mengelompokkan objek ke dalam kelompok-kelompok sesuai dengan karakteristik tertentu (Johnson & Wichern, 2007). Setiap pengamatan harus memiliki homogenitas yang tinggi dalam sebuah kelompok dan memiliki heterogenitas yang tinggi dengan kelompok yang lainnya. Analisis ini akan mengelompokkan objek sehingga setiap objek yang memiliki kesamaan dengan objek

lainnya akan berada dalam klaster yang sama (Sharma, 1996). *Cluster analysis* bertujuan untuk menentukan suatu kelompok yang alami dari kelompok-kelompok individu. Kelompok individu-individu ini bisa membentuk populasi yang lengkap atau suatu sampel dari populasi yang lebih besar. *Cluster analysis* akan mengalokasikan sekelompok individu pada suatu kelompok-kelompok yang saling bebas sehingga individu-individu di dalam kelompok itu mirip satu dengan yang lainnya, sementara itu individu-individu di dalam kelompok yang berbeda tidak mirip. Penyusunan kelompok ini biasa disebut dengan partisi (Ruswandi, 2008). Salah satu ukuran kemiripan yang dapat digunakan adalah jarak *euclidean* dan *mahalanobis*. Pada sudut pandang lainnya, beberapa metode *cluster* dapat dikelompokkan berdasarkan algoritma proses yang dilakukan, yaitu teknik yang berdasarkan ukuran jarak sebagai basis pengelompokannya.

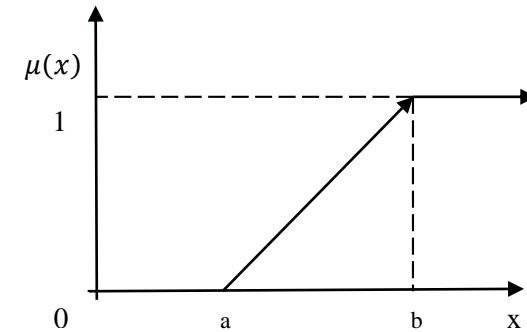
Metode berbasis ukuran jarak ini terdiri dari metode *cluster* berhierarki dengan penggabungan (*agglomerative*) dan pemisahan (*decisive*). Contohnya adalah metode pautan lengkap *complete linkage*, metode pautan rata-rata *average linkage*, metode Ward serta metode *cluster* tak berhierarki yaitu metode *k-means* (Andesberg, 1973). *Hierarchical clustering* maupun *k-means clustering* hanya memperhatikan ukuran jarak antar objek pengamatan tanpa mempertimbangkan aspek statistik lainnya, seperti distribusi data ataupun objek pengamatan pada *cluster* yang saling tumpang tindih. Sedangkan algoritma *fuzzy clustering* membentuk data kedalam struktur klaster yang berbentuk *spherical*, ellipsoid atau bentuk yang lain.

2.7 Fungsi Keanggotaan

Terdapat dua cara dalam melakukan pendekatan inisialisasi *membership* dalam himpunan sistem *fuzzy*, yaitu secara numerik dan fungsional. Pendekatan fungsional digunakan pada *Fuzzy clustering*. Kusumadewi dkk. (2004) menyatakan fungsi keanggotaan ini dapat dikatakan sebagai derajat keanggotaan (sebuah kurva) yang menunjukkan pemetaan titik obyek penelitian kedalam nilai keanggotaan yang berada dalam interval 0 hingga 1.

Representasi linier, kurva segitiga dan kurva trapesium merupakan fungsi keanggotaan yang paling sederhana dan dalam penelitian ini fungsi keanggotaan yang digunakan adalah representasi linier naik.

Fungsi keanggotaan representasi linier menunjukkan pemetaan input ke derajat keanggotaannya digambarkan sebagai sebuah garis lurus. Bentuk yang sangat sederhana ini cocok untuk sebuah himpunan data yang kurang memiliki kerangka konsep yang jelas. Contoh dari fungsi linier ditunjukkan pada Gambar 2.1, dimana fungsi tersebut menunjukkan kenaikan himpunan dimulai dari domain yang memiliki derajat keanggotaan nol [0] bergerak ke kanan menuju nilai *domain* yang memiliki derajat keanggotaan yang lebih tinggi. Gambar 2.1 menunjukkan representasi dari fungsi linear naik. Pada penelitian ini, derajat keanggotaan didasarkan pada hasil komponen skor faktor sedangkan nilai input berasal dari variabel realisasi pajak daerah. Derajat keanggotaan $\mu(x)$ dari fungsi linier naik pada Gambar 2.1 ditunjukkan pada Persamaan (2.5).



Gambar 2.1. Representasi Fungsi Linier Naik

$$\mu(x) = \begin{cases} 0 & ; x \leq a \\ \frac{(x-a)}{(b-a)} & ; a \leq x \leq b \\ 1 & ; x \geq b \end{cases} \quad (2.5)$$

Keterangan:

a = nilai domain yang mempunyai derajat keanggotaan nol

b = nilai domain yang mempunyai derajat keanggotaan satu
 x = nilai input yang akan diubah ke dalam bilangan *fuzzy*

Adapun nilai derajat keanggotaan dalam fungsi keanggotaan berasal dari komponen faktor f_1, f_2, \dots, f_N yang didapatkan oleh hasil analisis faktor. Nilai-nilai tersebut direpresentasikan melalui matriks U berukuran k -objek pada analisis kluster *fuzzy*.

2.8 Analisis Kluster Fuzzy K-Means

Metode kluster *k-means* merupakan salah satu metode kluster non hirarki dengan jumlah kluster sudah ditentukan sebelum pengelompokan yang dikembangkan Bezdek (1981). *Fuzzy k-means* merupakan salah satu metode pengelompokan yang dikembangkan dari *k-means* dengan menerapkan sifat *fuzzy* dalam fungsi keanggotaannya. Dalam metode ini dipergunakan variabel *membership function* yang merujuk pada seberapa besar kemungkinan suatu data bisa menjadi anggota ke dalam suatu kelompok. *Membership function* yang digunakan dapat berasal dari skor faktor. Konsep dasar *fuzzy k-means* berangkat dari ide mencari struktur kluster yang meminimumkan fungsi objektif yaitu jarak antar objek ke setiap titik pusat cluster dengan menggunakan Persamaan (2.6) dan Persamaan (2.7).

$$J_{FCM}(X, U, V) = \sum_{k=1}^L \sum_{j=1}^N (\mu_{ij})^m D_{ij} \quad (2.6)$$

$$J_{FCM}(X, U, V) = \sum_{k=1}^L \sum_{j=1}^N (\mu_{ij})^m \|v_k - f_j\| \quad (2.7)$$

dimana:

v_k = pusat kelompok ke- k ; $k = 1, 2, \dots, L$

μ_{ij} = derajat keanggotaan komponen faktor ke- j ; kelompok ke- i

m = bobot/*weighting exponent* ($m > 1$)

f_j = komponen faktor ke- j ; $j = 1, 2, \dots, N$.

Berdasarkan Bezdek (1981), langkah algoritma dari metode FKM sebagai berikut:

1. Memasukkan data yang akan dikelompokkan

2. Menentukan banyak jumlah kelompok yang akan dibentuk ($1 < c < N$), *weighting exponent* ($m > 1$), maksimum iterasi *error* terkecil yang diharapkan ($\varepsilon > 0$), fungsi objektif awal = 0, dan iterasi awal ($t = 1$),.
3. Membentuk matriks U sebagai elemen matriks partisi awal, seperti yang ditunjukkan pada Persamaan (2.8).

$$U = \begin{bmatrix} \mu_{11}(f_1) & \mu_{12}(f_2) & \cdots & \mu_{1N}(f_N) \\ \mu_{21}(f_1) & \mu_{22}(f_2) & \cdots & \mu_{2N}(f_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{c1}(f_1) & \mu_{c2}(f_2) & \cdots & \mu_{cN}(f_N) \end{bmatrix} \quad (2.8)$$

4. Menghitung pusat kelompok ke- i menurut Persamaan (2.9).

$$v_k = \frac{\sum_{j=1}^N (\mu_{ij})^m f_j}{\sum_{j=1}^N (\mu_{ij})^m}, k = 1, 2, \dots, L \quad (2.9)$$

5. Menghitung formula jarak *euclidean* menggunakan rumus pada Persamaan (2.10) dengan A merupakan matriks definit positif.

$$D_{ij} = \|x_j - v_k\| = (f_j - v_k)^T A (f_j - v_k) \quad (2.10)$$

6. Menghitung fungsi objektif yang ada di persamaan *Lagrange* (2.6) pada iterasi ke- t .
7. Menghitung nilai fungsi keanggotaan yang baru U_{t+1} sesuai dengan Persamaan (2.11).

$$\mu_{ij} = \left[\sum_{k=1}^L \left(\frac{D(f_j, v_k)}{D(f_j, v_j)} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (2.11)$$

8. Membandingkan nilai keanggotaan dalam matriks U hingga konvergen ketika ($\|U_{t+1} - U_t\| < \varepsilon$). Apabila $\|U_{t+1} - U_t\| \geq \varepsilon$ maka kembali ke langkah 4.

2.9 Analisis Kluster Fuzzy Gustafson-Kessel

Berdasarkan Balasko dkk. (2007), pengelompokan *fuzzy Gustafson-Kessel* merupakan pengembangan dari *fuzzy k-means* (FKM). Nilai pembentuk matriks pada metode pengelompokan ini disebut *adaptive distance norm* yang diperbarui di setiap iterasi. Sehingga, pengelompokan ini mampu lebih menyesuaikan bentuk geometris fungsi keanggotaan yang tepat untuk sebuah himpunan data.

Analisis kluster *fuzzy* menitikberatkan pada pengelompokan yang mempertimbangkan tingkat keanggotaan yang mencakup himpunan *fuzzy* sebagai landasan pembobotan. Perbedaan utama dari *fuzzy Gustafson-Kessel* dibandingkan dengan metode kluster hierarki maupun non hierarki ialah kemampuannya dalam menangani ketidakpastian. Mauliyadi dkk. (2003) menyatakan nilai akurasi analisis kluster *Gustafson-Kessel* pun lebih tinggi daripada analisis kluster *fuzzy c-means*. Amalia dkk. (2016) memperoleh nilai *icd-rate* FGK lebih rendah daripada FKM. Berdasarkan penelitian sebelumnya maka metode FGK lebih baik digunakan dibandingkan metode FKM.

Berdasarkan Balasko dkk. (2007), algoritma pengelompokan *fuzzy gustafson-kessel* adalah sebagai berikut.

1. Memasukkan data yang ingin dikelompokan.
2. Menentukan banyak kelompok yang akan dibentuk ($1 < c < N$), pembobot ($m > 1$), maksimum iterasi, error terkecil yang diharapkan ($\varepsilon > 0$), fungsi objektif awal = 0, dan iterasi awal ($t = 1$),.
3. Membentuk matriks U sebagai elemen matriks partisi awal sesuai dengan persamaan (2.8).
4. Menghitung pusat kelompok ke- k sesuai dengan persamaan (2.9).
5. Menghitung matriks kovarian kelompok F_k sesuai dengan Persamaan (2.12).

$$F_k = \frac{\sum_{j=1}^N (\mu_{ij})^m (f_j - v_k)^T A ((f_j - v_k))}{\sum_{j=1}^N (\mu_{ij})^m} \quad (2.12)$$

6. Menghitung jarak D_{ij} sesuai dengan persamaan (2.10) dengan $A = \left[\det(F_k)^{\frac{1}{N}} F_k^{-1} \right]$.
7. Menghitung fungsi objektif yang ada di persamaan (2.6) pada iterasi ke- t .
8. Menghitung nilai fungsi keanggotaan baru U_{t+1} sesuai dengan persamaan (2.11).
9. Membandingkan nilai keanggotaan dalam matriks U hingga konvergen ketika $\|U_{t+1} - U_t\| < \varepsilon$. Apabila $\|U_{t+1} - U_t\| \geq \varepsilon$ maka kembali ke langkah 4.

2.10 Calinski-Harabasz Pseudo F -Statistic

Penentuan banyaknya kelompok kluster optimum dilakukan dengan menghitung nilai *pseudo F -statistic* menggunakan rumus pada persamaan (2.13). Nilai *pseudo F -statistic* yang paling tinggi menunjukkan bahwa jumlah kelompok yang terbentuk telah optimal. Berikut rumus *pseudo F -statistic* menurut Orphin dkk. (2006).

$$Pseudo - F = \frac{\left(\frac{R^2}{c-1} \right)}{\left(\frac{1-R^2}{n-c} \right)} \quad (2.13)$$

dimana:

$$R^2 = \frac{(SST - SSE)}{SST} \quad (2.14)$$

$$SST = \sum_{g=1}^p \sum_{k=1}^L \sum_{j=1}^N (x_{gkj} - \bar{x}_j)^2 \quad (2.15)$$

$$SSE = \sum_{g=1}^p \sum_{k=1}^L \sum_{j=1}^N (x_{gkj} - \bar{x}_{kj})^2 \quad (2.16)$$

Keterangan:

x_{gkj} = objek ke- g pada kelompok ke- k dan komponen faktor ke- j

\bar{x}_j = rata-rata seluruh sampel pada komponen faktor j

\bar{x}_{kj} = rata-rata sampel kelompok ke- k dan komponen faktor ke- j

2.11 Evaluasi Kinerja Pengelompokan

Pengukuran kinerja hasil pengelompokan merupakan langkah untuk mengetahui validitas suatu pengelompokan. Berdasarkan Hair dkk. (2009) kelompok yang baik adalah memiliki kehomogenan yang tinggi antar anggota dalam kelompok dan keheterogenan yang tinggi antar kelompok. Berdasarkan Mingoti dkk. (2006) perbandingan dua atau lebih pengelompokan dapat diukur melalui tingkat rata-rata penyebaran dalam kelompok atau *internal cluster dispersion rate (icd-rate)*. Perbedaan keanggotaan tiap kelompok juga semakin rendah saat nilai *icd-rate* nya semakin rendah yang mengindikasikan metode tersebut lebih baik daripada nilai *icd-rate* yang lebih tinggi. Perhitungan *internal cluster dispersion rate* ditunjukkan pada Persamaan (2.17).

$$icdrate = 1 - \frac{(SST - SSE)}{SST} = 1 - \frac{SSB}{SST} = 1 - R^2 \quad (2.17)$$

2.12 Profil Kota Surabaya

Kota Surabaya merupakan salah satu kota metropolitan di Indonesia. Kota Surabaya merupakan ibukota provinsi di Jawa Timur dengan luas wilayah 326,36 km² dan memiliki jumlah penduduk sebanyak 2.848.583 jiwa di tahun 2015 (Badan Pusat Statistik, 2015). Wilayah kota Surabaya dibagi menjadi lima wilayah kerja administratif, yaitu Surabaya Pusat, Surabaya Timur, Surabaya Barat, Surabaya Utara, dan Surabaya Selatan (BPPK Surabaya, 2006).



Gambar 2.2. Peta Lahan Kota Surabaya

Tiap wilayah terdiri dari beberapa kecamatan dimana setiap kecamatan memiliki karakteristik potensi perpajakan yang berbeda-beda.

1. Surabaya Pusat

Wilayah ini terdiri dari kecamatan Tegalsari, Simokerto, Genteng, dan Bubutan. Di wilayah ini terdapat pusat bisnis Tunjungan dan kantor pemerintahan Kota Surabaya serta kantor pemerintahan Provinsi Jawa Timur. Potensi penerimaan pajak hotel, pajak reklame, pajak penerangan jalan dan pajak parkir sangat besar di area Surabaya pusat.

2. Surabaya Timur

Wilayah ini terdiri dari kecamatan Gubeng, Gunung Anyar, Sukolilo, Tambaksari, Mulyorejo, Rungkut, dan Tenggilis Mejoyo. Terdapat dua perguruan tinggi negeri terkemuka (ITS dan Universitas Airlangga) dan puluhan perguruan tinggi lainnya tersebar di Surabaya Timur ini secara tidak langsung membentuk sentra UMKM di sekitar kampus.

3. Surabaya Barat

Wilayah ini terdiri dari kecamatan Benowo, Pakal, Asem Rowo, Sukomanunggal, Tandes, Sambikerep, dan Lakarsantri. Terdapat perumahan elit Citraland yang membuat wilayah ini memiliki potensi penerimaan Pajak Bumi dan Bangunan (PBB) yang lebih tinggi daripada wilayah Surabaya lainnya.

4. Surabaya Utara

Wilayah ini terdiri dari kecamatan Bulak, Kenjeran, Semampir, Pabean Cantikan, dan Krembangan. Terdapat banyak sentra wisata seperti Pantai Ria Kenjeran, Wisata Religi Sunan Ampel serta pelabuhan Tanjung Perak yang ramai dengan aktivitas perdagangan.

5. Surabaya Selatan

Wilayah ini terdiri dari kecamatan Wonokromo, Wonocolo, Wiyung, Karang Pilang, Jambangan, Gayungan, Dukuh Pakis, dan Sawahan. Surabaya Industrial Estate Rungkut (SIER) terletak di Surabaya Selatan dan banyak perumahan berada di sekelilingnya.

2.13 Pengertian Pajak

Menurut Prof. Dr. M. J. H. Smeets, pajak adalah prestasi kepada pemerintah yang terutang melalui norma-norma umum dan yang dapat dipaksakan, tanpa adanya kontraprestasi yang dapat ditunjukkan dalam hal yang individual, dimaksud membiayai pengeluaran pemerintah (Brotodihardjo, 1993). Adapun fungsi pajak terbagi menjadi dua adalah sebagai berikut

1. Fungsi *Budgetair*

Pajak memberikan sumbangan terbesar dalam penerimaan atau pendapatan negara (60-70 persen penerimaan pajak memenuhi postur APBN).

2. Fungsi *Regulerend*

Pajak berfungsi sebagai alat untuk mengatur masyarakat atau melaksanakan kebijakan pemerintah dalam bidang menurut Halim dkk. (2014).

2.14 Pengertian Pajak Daerah

Menurut UU No. 28 Tahun 2009 pasal 1 (1) pajak daerah merupakan kontribusi wajib kepada daerah yang terutang oleh orang pribadi atau badan yang bersifat memaksa berdasarkan undang-undang, dengan tidak mendapat imbalan secara langsung dan digunakan untuk keperluan daerah bagi sebesar-besarnya kemakmuran rakyat. Pajak daerah digunakan untuk membiayai rumah tangga daerah. Berikut dasar hukum tujuh variabel pajak daerah yang digunakan dalam penelitian ini secara berurutan menurut UU No. 28 tahun 2009.

1. Pajak Hotel

Menurut UU No. 28 tahun 2009 pasal 1 (20) pajak hotel merupakan pajak atas pelayanan yang disediakan oleh hotel. Hotel merupakan fasilitas penyedia penginapan/peristirahatan termasuk jasa terkait lainnya dengan dipungut bayaran, yang mencakup juga motel, losmen, gubuk pariwisata, pesanggrahan, rumah penginapan dan sejenisnya, serta rumah kos dengan jumlah kamar lebih dari 10 (sepuluh).

2. Pajak Restoran

Menurut UU No. 28 tahun 2009 pasal 1 (22) pajak restoran merupakan pajak atas pelayanan yang disediakan oleh restoran. Restoran merupakan fasilitas penyedia makanan dan/atau minuman dengan dipungut bayaran, yang mencakup juga rumah makan, kafetaria, kantin, warung, bar, dan sejenisnya termasuk jasa boga/*catering*.

3. Pajak Hiburan

Menurut UU No. 28 tahun 2009 pasal 1 (24) pajak reklame adalah pajak penyelenggaraan hiburan. Hiburan merupakan semua jenis tontonan, pertunjukan, peramaian, dan/atau keramaian yang dinikmati dengan dipungut bayaran.

4. Pajak Penerangan Jalan

Menurut UU No. 28 tahun 2009 pasal 1 (28) pajak penerangan jalan adalah pajak atas penggunaan tenaga listrik, baik yang dihasilkan sendiri maupun diperoleh dari sumber lain.

5. Pajak Parkir

Menurut UU No. 28 tahun 2009 pasal 1 (31) pajak parkir merupakan pajak atas penyelenggaraan tempat parkir di luar badan jalan, baik yang disediakan berkaitan dengan pokok usaha maupun yang disediakan sebagai suatu usaha, termasuk penyediaan tempat penitipan kendaraan bermotor. Parkir adalah keadaan tidak bergerak suatu kendaraan yang tidak bersifat sementara.

6. Pajak Air Tanah

Menurut UU No. 28 tahun 2009 pasal 1 (33) pajak air tanah merupakan pajak atas pengambilan dan/atau pemanfaatan air tanah. Air tanah adalah air yang terdapat dalam lapisan tanah atau batuan dibawah permukaan tanah.

7. Pajak Bumi dan Bangunan

Menurut UU No. 28 tahun 2009 pasal 1 (37) pajak bumi dan bangunan merupakan pajak atas bumi dan/atau bangunan yang dimiliki, dikuasai, dan/atau dimanfaatkan oleh orang pribadi atau Badan, kecuali kawasan yang digunakan untuk kegiatan usaha perkebunan, perhutanan, dan pertambangan.

2.15 Studi Simulasi Pengelompokan

Simulasi merupakan suatu teknik meniru operasi-operasi atau proses-proses yang terjadi dalam suatu sistem dengan bantuan perangkat komputer dan dilandasi oleh beberapa asumsi tertentu sehingga sistem tersebut bisa dipelajari secara ilmiah (Law and Kelton, 1991). Dalam simulasi digunakan komputer untuk mempelajari sistem secara numerik, dimana dilakukan pengumpulan data untuk melakukan estimasi statistik karakteristik asli dari sistem. Simulasi merupakan alat yang tepat untuk digunakan terutama jika diharuskan untuk melakukan eksperimen dalam melakukan interpretasi komponen-komponen sistem, yang didalam penelitian ini adalah komponen sistem metode pengelompokan. Komponen akurasi atau kebaikan metode dalam membentuk klaster yang sesuai dengan keadaan sebenarnya diukur dalam studi simulasi

2.16 Penelitian Sebelumnya

Penelitian terdahulu terkait pajak daerah oleh Sirait (2015) menyimpulkan bahwa PDRB, jumlah hotel, jumlah pelanggan PLN, jumlah wisatawan dan kepadatan penduduk berpengaruh signifikan terhadap pajak daerah kota Surabaya. Selain itu, terdapat dua kelompok pajak daerah yang terbentuk dari penggunaan analisis klaster hierarki dengan metode *ward* (Lutfi, 2016) penggunaan metode *fuzzy c-shell* dan *fuzzy c-means* berdasarkan variabel pembentuk IPM (Widodo, 2012).

Riset terkait metode yang digunakan pada penelitian ini telah dilakukan oleh Amalia dkk. (2016) tentang Analisis Clustering Perusahaan Sub Sektor Perbankan berdasarkan Rasio Keuangan CAMELS Tahun 2014 menggunakan metode *fuzzy c-means* dan *fuzzy gustafson kessel* dengan hasil diperoleh kondisi optimum untuk kedua metode berdasarkan nilai Index XB sebanyak 2 *cluster* dengan metode terbaik adalah *fuzzy gustafson kessel* berdasarkan nilai icd-rate terkecil yaitu sebesar 0,7232. Berdasarkan jumlah *cluster* optimum dan metode terbaik, maka hasil pengujian MANOVA menghasilkan bahwa tidak ada perbedaan karakteristik antar cluster terhadap respon menurut Amalia dkk. (2016).

Penelitian terdahulu yang terkait fungsi keanggotaan oleh Kusuma (2017) menyimpulkan bahwa enam provinsi dari total sembilan provinsi di Sumatera menggunakan fungsi keanggotaan linier naik pada hasil pengelompokan *fuzzy gustafson-kessel*. Penelitian dilakukan Laraswati (2014) dan Alvionita (2017) yang bertujuan menentukan metode pautan jarak terbaik ini telah menggunakan kriteria nilai rasio simpangan baku terkecil dalam mendapatkan metode *agglomerative hierarchical clustering* terbaik dan klaster terbaik.

(Halaman ini sengaja dikosongkan)

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder dan data primer. Data sekunder diperoleh dari Badan Pengelolaan Keuangan dan Pajak Daerah (BPKPD) Kota Surabaya. Data yang diambil dari BPKPD yaitu data realisasi tujuh macam variabel pajak daerah di Kota Surabaya pada tahun 2016. Data primer diperoleh dari data bangkitan berdistribusi normal multivariat dan normal univariat yang akan dijelaskan di subbab 3.4. Struktur data penelitian yang digunakan dapat dilihat pada Tabel 3.1. Adapun unit penelitian yang digunakan pada penelitian ini dapat dilihat pada Tabel 3.2.

Tabel 3.1 Struktur Data Penelitian

Kecamatan ke- i	X_1	X_2	X_3	...	X_7
1	$X_{1,1}$	$X_{2,1}$	$X_{3,1}$...	$X_{7,1}$
2	$X_{1,2}$	$X_{2,2}$	$X_{3,2}$...	$X_{7,2}$
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
31	$X_{1,31}$	$X_{2,31}$	$X_{3,31}$...	$X_{7,31}$

Tabel 3.2 Unit Penelitian

No	Kecamatan	No	Kecamatan	No	Kecamatan
1	Asemrowo	12	Kenjeran	23	Sukolilo
2	Benowo	13	Krembangan	24	Sukomanunggal
3	Bubutan	14	Lakarsantri	25	Tambaksari
4	Bulak	15	Mulyorejo	26	Tandes
5	Dukuh Pakis	16	Pabean Cantian	27	Tegalsari
6	Gayungan	17	Pakal	28	Tenggiling Mejoyo
7	Genteng	18	Rungkut	29	Wiyung
8	Gubeng	19	Sambikerep	30	Wonocolo
9	Gunung Anyar	20	Sawahan	31	Wonokromo
10	Jambangan	21	Semampir		
11	Karang Pilang	22	Simokerto		

3.2 Variabel Penelitian

Berdasarkan Peraturan Daerah Kota Surabaya Nomor 4 Tahun 2011 tentang pajak daerah, maka pada penelitian ini akan digunakan variabel seluruhnya berskala rasio pada Tabel 3.3.

Tabel 3.3 Variabel Penelitian

Variabel	Nama Variabel
X_1	Pajak Air Tanah
X_2	Pajak Hiburan
X_3	Pajak Penerangan Jalan
X_4	Pajak Hotel
X_5	Pajak Restoran
X_6	Pajak Parkir
X_7	Pajak Bumi dan Bangunan

3.3 Definisi Operasional

Adapun penjelasan tentang variabel variabel-variabel penelitian akan dijelaskan pada definisi operasional sebagai berikut.

a. Pajak Air Tanah

Pajak air tanah merupakan pajak atas pengambilan dan/atau pemanfaatan air tanah. Air tanah adalah air yang terdapat dalam lapisan tanah atau batuan dibawah permukaan tanah. Objek pajak air tanah adalah pengambilan dan/atau pemanfaatan air tanah dengan pengecualian pengambilan dan/atau pemanfaatan air tanah untuk keperluan dasar rumah tangga, pengairan pertanian dan perikanan rakyat, serta peribadatan.

b. Pajak Hiburan

Pajak hiburan adalah pajak penyelenggaraan hiburan. Hiburan merupakan semua jenis tontonan, pertunjukan, peramaian, dan/atau keramaian yang dinikmati dengan dipungut bayaran. Objek pajak hiburan adalah jasa penyelenggaraan hiburan dengan dipungut bayaran, meliputi: tontonan film, pagelaran kesenian, musik, tari, kontes kecantikan, pameran, diskotik, karaoke, klab malam, sirkus, akrobat, sulap, permainan bilyar, golf, boling, pacuan kuda, panti pijat dan pertandingan olahraga.

c. Pajak Penerangan Jalan

Pajak penerangan jalan adalah pajak atas penggunaan tenaga listrik, baik yang dihasilkan sendiri maupun diperoleh dari sumber lain. Objek pajak penerangan jalan adalah penggunaan

tenaga listrik, baik yang dihasilkan sendiri maupun yang diperoleh dari sumber lain dari PLN/bukan PLN.

d. Pajak Hotel

Pajak hotel merupakan pajak atas pelayanan yang disediakan oleh hotel. Hotel merupakan fasilitas penyedia jasa penginapan/peristirahatan termasuk jasa terkait lainnya dengan dipungut bayaran, yang mencakup juga motel, losmen, gubuk pariwisata, pesanggrahan, rumah penginapan dan sejenisnya, serta rumah kos dengan jumlah kamar lebih dari 10 (sepuluh). Objek pajak hotel adalah pelayanan yang disediakan oleh hotel dengan pembayaran, termasuk jasa penunjang sebagai kelengkapan hotel yang sifatnya memberikan kemudahan dan kenyamanan, termasuk fasilitas olahraga dan hiburan.

e. Pajak Restoran

Pajak restoran merupakan pajak atas pelayanan yang disediakan oleh restoran. Restoran merupakan fasilitas penyedia makanan dan / atau minuman dengan dipungut bayaran, yang mencakup juga rumah makan, kafetaria, kantin, warung, bar, dan sejenisnya termasuk jasa boga/*catering*. Objek pajak restoran adalah pelayanan yang disediakan oleh restoran.

f. Pajak Parkir

Pajak parkir merupakan pajak atas penyelenggaraan tempat parkir di luar badan jalan, baik yang disediakan berkaitan dengan pokok usaha maupun disediakan sebagai suatu usaha, termasuk penyediaan tempat penitipan kendaraan bermotor. Parkir adalah keadaan tidak bergerak suatu kendaraan yang tidak bersifat sementara. Objek Pajak Parkir adalah penyelenggaraan tempat Parkir di luar badan jalan, baik yang disediakan berkaitan dengan pokok usaha maupun yang disediakan sebagai suatu usaha, termasuk penyediaan tempat penitipan kendaraan bermotor.

g. Pajak Bumi dan Bangunan

Pajak bumi dan bangunan merupakan pajak atas bumi dan/atau bangunan yang dimiliki, dikuasai, dan/atau dimanfaatkan oleh

orang pribadi atau Badan, kecuali kawasan yang digunakan untuk kegiatan usaha perkebunan, perhutanan, dan pertambangan. Objek pajak bumi dan bangunan perdesaan dan perkotaan adalah bumi dan/atau bangunan yang dimiliki, dikuasai, dan/atau dimanfaatkan oleh orang pribadi atau badan, kecuali kawasan yang digunakan untuk kegiatan usaha perkebunan, perhutanan, dan pertambangan.

3.4 Desain Studi Simulasi Pengelompokan

Berikut ini adalah langkah yang digunakan dalam melakukan studi simulasi pengelompokan:

1. Menentukan tujuan simulasi yaitu menunjukkan adanya perbedaan akurasi kedua metode *clustering* pada saat melakukan pengelompokan.
2. Menentukan parameter tujuh dataset bangkitan data berdistribusi normal multivariat dan tujuh dataset bangkitan data berdistribusi normal univariat dengan beberapa jumlah pengamatan berurutan dari yang terendah yaitu 30, 60, 90, 120, 150, 300 dan 1500 pengamatan dimana seluruh pembangkitan data berdistribusi normal multivariat memiliki matriks varian kovarian $S = \begin{bmatrix} 1,0 & 0,7 \\ 0,7 & 1,0 \end{bmatrix}$ dan $\sigma_{15} = 2$ serta $\sigma_{16} = \sigma_{17} = \dots = \sigma_{21} = 20$.
3. Membentuk fungsi keanggotaan linier naik u_1 dan u_2 untuk tiap c pengamatan yaitu $c = 30, c = 60, c = 90, c = 120, c = 150, c = 300, c = 1500$. Setiap objek pengamatan diharuskan memiliki derajat keanggotaan yang menyesuaikan bentuk fungsi keanggotaan linier naik
4. Membangkitkan data berdasarkan langkah no. 2.

Untuk $\Sigma = \begin{bmatrix} 1,0 & 0,7 \\ 0,7 & 1,0 \end{bmatrix}$ pada bangkitan normal multivariat maka:

$$c = 30, (X_1, X_2) \sim N_2 \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma \right)$$

$$\mu_1 = \begin{bmatrix} E(X_{1[1:15]}) \\ E(X_{1[16:30]}) \end{bmatrix} = \begin{bmatrix} 9 \\ 9 \end{bmatrix}, \mu_2 = \begin{bmatrix} E(X_{2[1:15]}) \\ E(X_{2[16:30]}) \end{bmatrix} = \begin{bmatrix} 9 \\ 8,9 \end{bmatrix}$$

$$c = 60, (X_3, X_4) \sim N_2 \left(\begin{bmatrix} \mu_3 \\ \mu_4 \end{bmatrix}, \Sigma \right)$$

$$\mu_3 = \begin{bmatrix} E(X_{3[1:30]}) \\ E(X_{3[31:60]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 90 \end{bmatrix}, \mu_4 = \begin{bmatrix} E(X_{4[1:30]}) \\ E(X_{4[31:60]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}$$

$$c = 90, (X_5, X_6) \sim N_2 \left(\begin{bmatrix} \mu_5 \\ \mu_6 \end{bmatrix}, \Sigma \right)$$

$$\mu_5 = \begin{bmatrix} E(X_{5[1:45]}) \\ E(X_{5[46:90]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 90 \end{bmatrix}, \mu_6 = \begin{bmatrix} E(X_{6[1:45]}) \\ E(X_{6[46:90]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}$$

$$c = 120, (X_7, X_8) \sim N_2 \left(\begin{bmatrix} \mu_7 \\ \mu_8 \end{bmatrix}, \Sigma \right)$$

$$\mu_7 = \begin{bmatrix} E(X_{7[1:60]}) \\ E(X_{7[61:120]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 90 \end{bmatrix}, \mu_8 = \begin{bmatrix} E(X_{8[1:60]}) \\ E(X_{8[61:120]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}$$

$$c = 150, (X_9, X_{10}) \sim N_2 \left(\begin{bmatrix} \mu_9 \\ \mu_{10} \end{bmatrix}, \Sigma \right)$$

$$\mu_9 = \begin{bmatrix} E(X_{9[1:75]}) \\ E(X_{9[76:150]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 90 \end{bmatrix}, \mu_{10} = \begin{bmatrix} E(X_{10[1:75]}) \\ E(X_{10[76:150]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}$$

$$c = 300, (X_{11}, X_{12}) \sim N_2 \left(\begin{bmatrix} \mu_{11} \\ \mu_{12} \end{bmatrix}, \Sigma \right)$$

$$\mu_{11} = \begin{bmatrix} E(X_{11[1:150]}) \\ E(X_{11[151:300]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 90 \end{bmatrix}, \mu_{12} = \begin{bmatrix} E(X_{12[1:75]}) \\ E(X_{12[76:150]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}$$

$$c = 1500, (X_{13}, X_{14}) \sim N_2 \left(\begin{bmatrix} \mu_{13} \\ \mu_{14} \end{bmatrix}, \Sigma \right)$$

$$\mu_{13} = \begin{bmatrix} E(X_{13[1:750]}) \\ E(X_{13[751:1500]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 90 \end{bmatrix}, \mu_{14} = \begin{bmatrix} E(X_{14[1:750]}) \\ E(X_{14[751:1500]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}$$

Sedangkan pada distribusi normal univariat sebagai berikut:

$$c = 30, \sigma_{15} = \sigma_{16} = 2$$

$$X_{15} \sim N(\mu_{15}, \sigma^2), X_{16} \sim N(\mu_{16}, \sigma^2)$$

$$\mu_{15} = \begin{bmatrix} E(X_{15[1:15]}) \\ E(X_{15[16:30]}) \end{bmatrix} = \begin{bmatrix} 9,0 \\ 8,9 \end{bmatrix}, \mu_{16} = \begin{bmatrix} E(X_{16[1:15]}) \\ E(X_{16[16:30]}) \end{bmatrix} = \begin{bmatrix} 8,9 \\ 90 \end{bmatrix}$$

$$c = 60, \sigma_{17} = \sigma_{18} = 20$$

$$X_{17} \sim N(\mu_{17}, \sigma^2), X_{18} \sim N(\mu_{18}, \sigma^2)$$

$$\mu_{17} = \begin{bmatrix} E(X_{17[1:30]}) \\ E(X_{17[31:60]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}, \mu_{18} = \begin{bmatrix} E(X_{18[1:30]}) \\ E(X_{18[31:60]}) \end{bmatrix} = \begin{bmatrix} 89 \\ 90 \end{bmatrix}$$

$$c = 90, \sigma_{19} = \sigma_{20} = 20$$

$$X_{19} \sim N(\mu_{19}, \sigma^2), X_{20} \sim N(\mu_{20}, \sigma^2)$$

$$\mu_{19} = \begin{bmatrix} E(X_{19[1:45]}) \\ E(X_{19[46:90]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}, \mu_{20} = \begin{bmatrix} E(X_{20[1:45]}) \\ E(X_{20[46:90]}) \end{bmatrix} = \begin{bmatrix} 89 \\ 90 \end{bmatrix}$$

$$c = 120, \sigma_{21} = \sigma_{22} = 20$$

$$X_{21} \sim N(\mu_{21}, \sigma^2), X_{22} \sim N(\mu_{22}, \sigma^2)$$

$$\mu_{21} = \begin{bmatrix} E(X_{21[1:60]}) \\ E(X_{21[61:120]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}, \mu_{22} = \begin{bmatrix} E(X_{22[1:60]}) \\ E(X_{22[61:120]}) \end{bmatrix} = \begin{bmatrix} 89 \\ 90 \end{bmatrix}$$

$$c = 150, \sigma_{23} = \sigma_{24} = 20$$

$$X_{23} \sim N(\mu_{23}, \sigma^2), X_{24} \sim N(\mu_{24}, \sigma^2)$$

$$\mu_{23} = \begin{bmatrix} E(X_{23[1:75]}) \\ E(X_{23[76:150]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}, \mu_{24} = \begin{bmatrix} E(X_{24[1:75]}) \\ E(X_{24[76:150]}) \end{bmatrix} = \begin{bmatrix} 89 \\ 90 \end{bmatrix}$$

$$c = 300, \sigma_{25} = \sigma_{26} = 20$$

$$X_{25} \sim N(\mu_{25}, \sigma^2), X_{26} \sim N(\mu_{26}, \sigma^2)$$

$$\mu_{25} = \begin{bmatrix} E(X_{25[1:150]}) \\ E(X_{25[151:300]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}, \mu_{26} = \begin{bmatrix} E(X_{26[1:150]}) \\ E(X_{26[151:300]}) \end{bmatrix} = \begin{bmatrix} 89 \\ 90 \end{bmatrix}$$

$$c = 1500, \sigma_{27} = \sigma_{28} = 20$$

$$X_{27} \sim N(\mu_{27}, \sigma^2), X_{28} \sim N(\mu_{28}, \sigma^2)$$

$$\mu_{27} = \begin{bmatrix} E(X_{27[1:750]}) \\ E(X_{27[751:1500]}) \end{bmatrix} = \begin{bmatrix} 90 \\ 89 \end{bmatrix}, \mu_{28} = \begin{bmatrix} E(X_{28[1:750]}) \\ E(X_{28[751:1500]}) \end{bmatrix} = \begin{bmatrix} 89 \\ 90 \end{bmatrix}$$

5. Melakukan pengelompokan dengan kedua metode *clustering* pada setiap dataset dengan melakukan perulangan sebanyak sepuluh kali untuk memperoleh keyakinan untuk mendekati kebenaran tingkat akurasi metode.
6. Menginterpretasi hasil akurasi dengan grafik *scatterplot* dan tabel perbandingan.

3.5 Langkah Analisis

Berikut ini adalah langkah analisis yang digunakan dalam melakukan penelitian :

1. Melakukan studi simulasi pengelompokan untuk memperoleh tingkat akurasi kedua metode *clustering* pada data bangkitan berdistribusi normal multivariat dan normal univariat.
2. Mendeskripsikan karakteristik kecamatan berdasarkan realisasi pajak daerah di Kota Surabaya tahun 2016 menggunakan statistika deskriptif.
 - a. Penentuan input variabel pajak daerah yang digunakan.
 - b. Mengeksplorasi input variabel pajak daerah dengan statistika deskriptif.
3. Melakukan pengelompokan kecamatan berdasarkan realisasi pajak daerah di Kota Surabaya tahun 2016. Adapun langkah-langkah analisis adalah sebagai berikut :
 - a. Menguji asumsi distribusi normal multivariat.
 - b. Menguji kecukupan data dengan uji KMO.
 - c. Menguji korelasi antar variabel dengan uji independensi.
 - d. Mengestimasi *loading factor* dengan metode analisis komponen utama.

Mengklasterkan kecamatan berdasarkan *loading factor* yang terbentuk, menggunakan metode *fuzzy k-means clustering* dan *fuzzy gustafson-kessel clustering* dengan jumlah klaster sebanyak 2 sampai dengan 4.

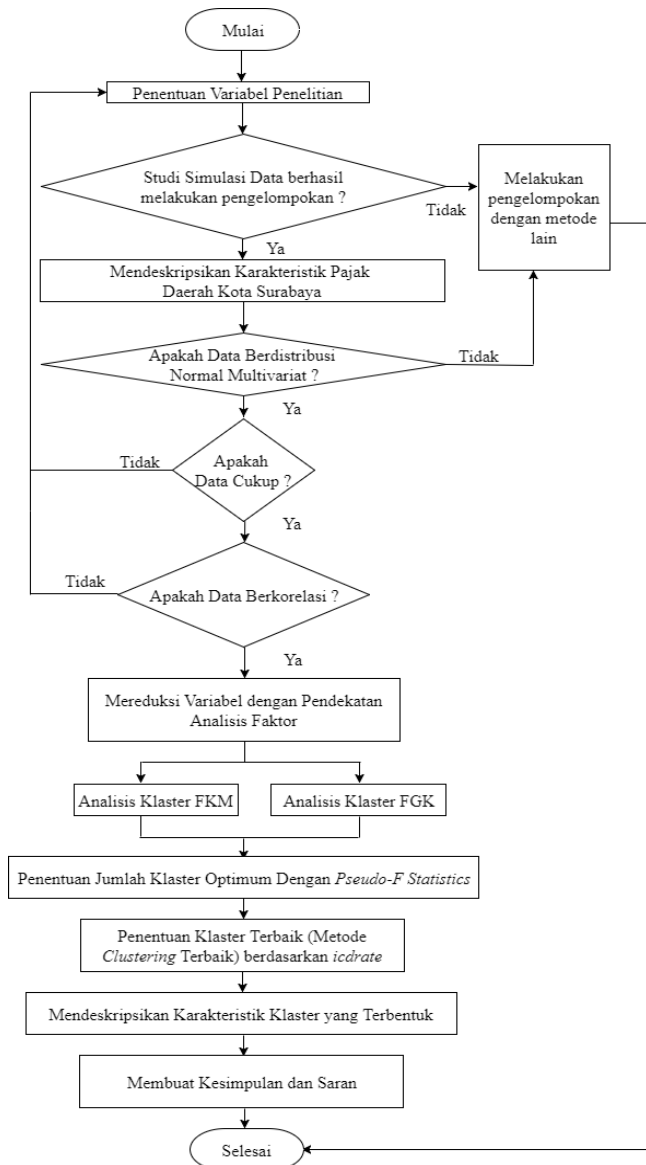
Adapun langkah-langkah analisis adalah sebagai berikut :

- a. Menentukan derajat keanggotaan melalui fungsi keanggotaan sebagai representasi metode FKM dan FGK

dengan pereduksian variabel menggunakan pendekatan analisis komponen utama.

- b. Melakukan pemilihan variabel yang mewakili setiap faktor komponen untuk pengelompokan berdasarkan *loading factor* yang terbentuk.
 - c. Melakukan analisis *clustering* menggunakan metode FKM dengan input berupa data realisasi pajak daerah dan inisiasi *membership function* linier naik dua hingga empat klaster dengan bantuan *fanny package*.
 - d. Melakukan analisis *clustering* menggunakan metode FGK dengan input berupa data realisasi pajak daerah dan inisiasi *membership function* linier naik dua hingga empat klaster dengan bantuan *fclust package*.
 - e. Menentukan jumlah klaster optimum menggunakan *Pseudo-F Statistics*.
 - f. Melakukan evaluasi kinerja metode pengelompokan *fuzzy k-means clustering* dan *fuzzy gustafson-kessel clustering* menggunakan kriteria *icd-rate*.
4. Melakukan analisis statistika deskriptif untuk mengetahui karakteristik masing-masing klaster yang terbentuk berdasarkan hasil pengelompokan.

Langkah-langkah analisis secara umum dapat disajikan pada diagram alir sebagai berikut.



Gambar 3.1 Diagram Alir Penelitian

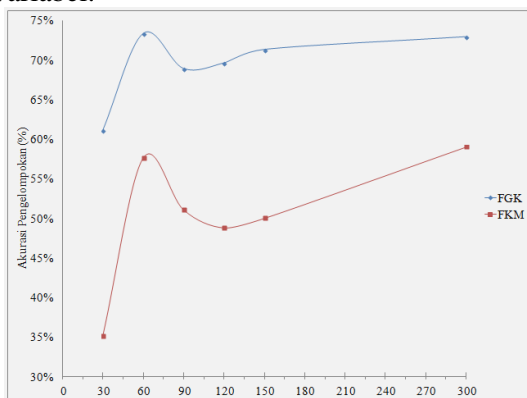
(Halaman ini sengaja dikosongkan)

BAB IV ANALISIS DAN PEMBAHASAN

Pada bab analisis dan pembahasan ini dilakukan analisis untuk menjawab permasalahan dari penelitian mengenai karakteristik, pengelompokan pajak daerah di kota Surabaya berdasarkan tujuh variabel yang telah ditentukan dengan metode *fuzzy k-means* (FKM) *clustering* dan *fuzzy Gustafson-Kessel* (FGK) *clustering*.

4.1 Studi Simulasi Pengelompokan Kecamatan

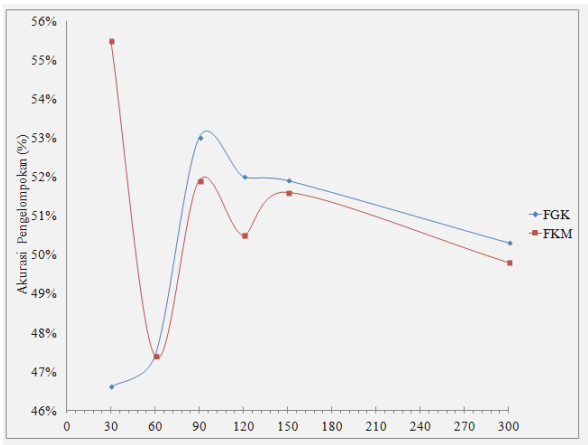
Studi simulasi pengelompokan pada penelitian ini bertujuan untuk menunjukkan tingkat keberhasilan suatu metode dalam melakukan pengelompokan terhadap populasi. Simulasi pengelompokan dimulai dengan pembangkitan gugus data berdistribusi normal multivariat dan normal univariat sebanyak 14 gugus data yang setiap gugus data dibangkitkan sebanyak 10 kali dengan 2 variabel.



Gambar 4.1. Akurasi FGK-FKM Normal Multivariat K=2

Pembangkitan data berdistribusi normal multivariat dilakukan sesuai dengan asumsi data untuk kedua metode *clustering* yang harus berdistribusi normal multivariat dengan korelasi yang kuat antar variabel. Korelasi antar kedua variabel telah diatur sebesar 0,7. Berdasarkan Gambar 4.1, akurasi metode FGK selalu

berada diatas akurasi metode FKM yaitu antara 61% hingga 74% pada $c = 300$.



Gambar 4.2. Akurasi FGK-FKM Normal Univariat K=2

Simulasi pembangkitan data berdistribusi normal univariat dilakukan untuk mengakomodasi *case* yang tidak mampu memenuhi asumsi data berdistribusi normal multivariat. Berdasarkan Gambar 4.2, akurasi metode FKM lebih baik pada dataset kecil $c = 30$ namun setelah itu terjadi *overlapping* dimana akurasi metode FGK selalu mengungguli metode FKM pada $c = 60$ hingga $c = 300$.

Tabel 4.1 Perbandingan Akurasi Normal Multivariat

Output ke-	FGK c=300	FKM c=300	FGK c=1500	FKM c=1500
1	0,730	0,390	0,778	0,630
2	0,790	0,646	0,746	0,625
3	0,703	0,610	0,756	0,601
4	0,700	0,610	0,709	0,382
5	0,700	0,610	0,758	0,617
6	0,700	0,610	0,754	0,372
7	0,733	0,666	0,748	0,650
8	0,773	0,583	0,731	0,596
9	0,733	0,593	0,768	0,612
10	0,723	0,646	0,754	0,61
Rerata	0,729	0,590	0,749	0,565

Berdasarkan Tabel 4.1, terjadi perbedaan arah pergerakan akurasi antar kedua metode dalam mengelompokkan data berdistribusi normal multivariat ketika dataset diperbesar. Kenaikan akurasi sebesar 0,02% terjadi pada metode FGK sedangkan penurunan sebesar 0,025% pada metode FKM. Terdapat perbedaan akurasi yang mencolok diantara kedua metode. Akurasi FGK yang selalu diatas 70% mengindikasikan bahwa metode FGK lebih baik dari metode FKM dalam aspek kesamaan pengelompokan suatu objek pengamatan yang sesuai dengan kondisi sebenarnya. Hasil studi simulasi secara lengkap dapat dilihat pada Lampiran 1 dan Lampiran 6 sampai Lampiran 9.

4.2 Karakteristik Pajak Daerah di Kota Surabaya

Kota Surabaya merupakan salah satu kota dengan penyumbang pajak daerah terbesar di Indonesia yang memiliki 31 kecamatan. Statistika deskriptif perlu dilakukan untuk mengetahui karakteristik data berdasarkan realisasi pajak daerah seluruh kecamatan di Kota Surabaya.

Deskripsi kondisi kecamatan di Kota Surabaya dari tujuh variabel dalam menentukan hasil pengelompokan menggunakan *fuzzy k-means clustering* dan *fuzzy Gustafson-Kessel clustering* dapat digambarkan dengan analisis statistika deskriptif. Berikut adalah karakteristik data seluruh kecamatan di Kota Surabaya berdasarkan tujuh variabel yang telah ditentukan.

Tabel 4.2 Karakteristik Pajak Daerah Surabaya (dalam miliar rupiah)

Variabel	Rata-Rata	Minimum	Maksimum
X_1	0,047	0,012	0,016
X_2	1,931	0	10,048
X_3	11,501	0	187,549
X_4	6,800	0	58,262
X_5	10,508	0,002	53,217
X_6	2,122	0,007	10,449
X_7	27,403	5,332	66,432

Tabel 4.2 menunjukkan variabel-variabel tersebut dibedakan menjadi tujuh variabel yaitu pajak air tanah, pajak hiburan, pajak penerangan jalan, pajak hotel, pajak restoran, pajak parkir serta pajak bumi dan bangunan dalam miliar rupiah. Variabel pajak air tanah memiliki nilai rata-rata sebesar 47,277 juta rupiah dengan

nilai maksimum sebesar 116,717 juta rupiah pada kecamatan Gubeng dan nilai minimum sebesar 12,905 juta rupiah pada kecamatan Jambangan. Variabel pajak hiburan memiliki nilai rata-rata sebesar 1,931 miliar rupiah dengan nilai maksimum sebesar 10,048 miliar rupiah pada kecamatan Genteng dan nilai minimum sebesar 0 rupiah pada kecamatan Benowo, Kenjeran dan Semampir. Variabel pajak penerangan jalan memiliki nilai rata-rata sebesar 11,501 miliar rupiah dengan nilai maksimum sebesar 187,549 miliar rupiah pada kecamatan Gubeng dan nilai minimum sebesar 0 rupiah pada kecamatan Benowo, Kenjeran dan Pakal.

Variabel pajak hotel memiliki nilai rata-rata sebesar 6,8 miliar rupiah dengan nilai maksimum sebesar 58,262 miliar rupiah pada kecamatan Tegalsari dan nilai minimum sebesar 0 rupiah pada kecamatan Benowo, Gunung Anyar, Kenjeran dan Pakal. Variabel pajak restoran memiliki nilai rata-rata sebesar 10,508 miliar Rupiah dengan nilai maksimum sebesar 53,217 miliar rupiah pada kecamatan Tegalsari dan nilai minimum sebesar 2,703 juta rupiah pada kecamatan Pakal. Variabel pajak parkir memiliki nilai rata-rata sebesar 2,122 miliar rupiah dengan nilai maksimum sebesar 10,449 miliar rupiah pada kecamatan Genteng dan nilai minimum sebesar 7,628 juta rupiah pada kecamatan Pakal. Variabel pajak bumi dan bangunan (PBB) memiliki nilai rata-rata sebesar 27,403 miliar rupiah dengan nilai maksimum sebesar 66,432 miliar rupiah pada kecamatan Mulyorejo dan nilai minimum sebesar 5,332 miliar rupiah pada kecamatan Jambangan.

4.3 Pra-Pengelompokan Kecamatan di Kota Surabaya

Setelah studi simulasi data pengelompokan dilakukan untuk pengecekan tingkat validasi kedua metode *clustering* dalam melakukan pengelompokan, penentuan hasil kelompok realisasi pajak daerah kecamatan di Kota Surabaya dengan menggunakan *fuzzy k-means clustering* dan *fuzzy Gustafson-Kessel clustering* maka akan dipilih beberapa variabel yang dapat mewakili pengelompokan kecamatan menggunakan analisis faktor.

Pemenuhan asumsi distribusi normal multivariat, asumsi kecukupan data dan asumsi korelasi harus dilakukan terlebih dahulu sebelum penggunaan analisis faktor yang dilanjutkan oleh pengelompokan dengan FKM & FGK beserta evaluasi kinerja metode FKM & FGK.

a. Pemenuhan Asumsi Distribusi Normal Multivariat

Pengujian distribusi normal multivariat adalah syarat dalam analisis faktor dengan pendekatan *principal component analysis* dan analisis kluster non hirarki menggunakan metode *k-means*. Berdasarkan hasil pengujian melalui nilai koefisien korelasi *pearson* terhadap data tujuh variabel pajak daerah ini didapatkan nilai koefisien korelasi *pearson* sebesar 0,962 dengan taraf signifikan 5% dan jumlah data sebanyak 31 didapatkan nilai r_{tabel} sebesar 0,3550 dengan derajat bebas 29 sehingga statistik uji lebih besar daripada nilai r_{tabel} yang berarti bahwa data pajak daerah 31 kecamatan di Kota Surabaya telah berdistribusi normal multivariat.

b. Pemenuhan Asumsi Kecukupan Data

Pengujian asumsi kecukupan data ini dilakukan untuk mengetahui kelayakan data untuk dianalisis menggunakan analisis faktor. Apabila nilai KMO (*Kaiser-Meyer-Olkin*) lebih besar dari 0,50 maka variabel penelitian telah layak untuk dianalisis. Berdasarkan hasil pengujian melalui nilai KMO (*Kaiser-Meyer-Olkin*) terhadap data pajak daerah Kota Surabaya ini didapatkan nilai KMO sebesar 0,67 sehingga statistik uji lebih besar daripada nilai 0,50 yang berarti bahwa seluruh data pajak daerah 31 kecamatan di Kota Surabaya layak dianalisis menggunakan analisis faktor.

c. Pemenuhan Asumsi Korelasi

Asumsi korelasi harus dipenuhi untuk melakukan analisis faktor setelah data telah memenuhi syarat kelayakan (kecukupan) data. Salah satu metode untuk mengetahui adanya korelasi antar variabel dapat menggunakan uji independensi *bartlett*. Hipotesis uji ini adalah sebagai berikut.

Hipotesis

$H_0 : \rho = \mathbf{I}$ (tidak terdapat korelasi)

$H_1 : \rho \neq \mathbf{I}$ (ada korelasi)

Berdasarkan hasil pengujian independensi terhadap data pajak daerah Kota Surabaya ini didapatkan nilai *Chi-Square* sebesar 147,954 dengan *P-value* kurang dari $\alpha = 0,05$ sehingga keputusannya adalah tolak H_0 yang menunjukkan adanya korelasi antar tujuh variabel pajak daerah. Pemenuhan asumsi distribusi normal multivariat, kecukupan data dan korelasi telah dilakukan sehingga analisis dapat dilanjutkan ke analisis faktor.

Berdasarkan hasil uji kecukupan data dan uji independensi dapat disimpulkan bahwa data layak (cukup) untuk dilakukan analisis faktor dan terdapat korelasi antar variabel. *Communalities* yang terdapat pada Tabel 4.3 menjelaskan metode ekstraksi komponen utama yang bertujuan untuk mengetahui seberapa besar masing-masing variabel berperan dalam menjelaskan faktor yang terbentuk. Penentuan jumlah faktor yang terbentuk dapat dilakukan secara matematis dan *scree plot* (Lampiran 4).

Penentuan jumlah faktor yang terbentuk secara matematis dapat dilakukan dengan perhitungan *eigenvalue* yang lebih dari 1 pada Tabel 4.4 atau juga dapat dilihat dari proporsi keragaman data dalam menjelaskan total varians sampel yang lebih dari 70%. Susunan *eigenvalue* selalu diurutkan dari yang terbesar sampai yang terkecil, dengan kriteria bahwa angka *eigenvalue* lebih kecil dari 1 tidak digunakan dalam menghitung jumlah faktor yang terbentuk.

Tabel 4.3 *Communalities* Data seluruh Kecamatan di Kota Surabaya

Variabel	Initial	Extraction
X_1	1,000	0,667
X_2	1,000	0,916
X_3	1,000	0,754
X_4	1,000	0,713
X_5	1,000	0,908
X_6	1,000	0,769
X_7	1,000	0,448

Variabel pajak hiburan memiliki peran yang paling besar dalam menjelaskan faktor yang terbentuk, yaitu sebesar 91%, diikuti variabel pajak restoran sebesar 90%, pajak parkir sebesar

76%, pajak penerangan jalan sebesar 75%, pajak hotel sebesar 71% dan pajak air tanah sebesar 66%. Variabel pajak bumi dan bangunan memiliki peran paling kecil dalam menjelaskan faktor yang terbentuk, yaitu sebesar 44%.

Tabel 4.4 Proporsi Keragaman Seluruh Kecamatan di Kota Surabaya

Komponen	Total	% Varians	Kumulatif %
1	3,904	55,772	55,772
2	1,271	18,151	73,923
3	0,767	10,961	84,884
4	0,511	7,296	92,181
5	0,372	5,313	97,494
6	0,134	1,910	99,404
7	0,042	0,596	100,000

Berdasarkan Tabel 4.4, pada mulanya terdapat 7 komponen atau faktor sesuai dengan jumlah variabel asli. Selanjutnya dengan melihat total *eigenvalue*, hanya faktor ke-1 dan faktor ke-2 yang memiliki total *eigenvalue* lebih dari 1 sedangkan faktor ke-3 memiliki nilai total *eigenvalue* kurang dari 1, sehingga proses pembentukan faktor berhenti pada faktor ke-2 yang berarti ada 2 faktor baru yang akan terbentuk. Prosentase varians kedua faktor dalam menjelaskan seluruh variabel asli adalah faktor ke-1 sebesar 55,772% dan faktor ke-2 sebesar 18,151%, sehingga total varians yang dapat dijelaskan oleh ketiga faktor terhadap variabel asli adalah sebesar 73,923%. Dalam analisis secara grafik maupun secara matematis, menghasilkan nilai yang sama yaitu ada dua komponen faktor baru yang akan terbentuk yang mewakili tujuh variabel yang tersedia untuk melakukan pengelompokan.

4.4 Pengelompokan Kecamatan di Kota Surabaya

Dalam melakukan pengelompokan kecamatan di Kota Surabaya terdapat koefisien matriks korelasi yang digunakan dalam menentukan variabel yang terpilih untuk mewakili masing-masing komponen faktor. Proses penentuan variabel yang menjadi anggota dari masing-masing komponen faktor dilakukan dengan membandingkan besarnya korelasi antara suatu variabel dengan masing-masing faktor yang terbentuk. Suatu variabel x_i menjadi anggota dari suatu faktor f_i , apabila variabel x_i memiliki korelasi paling kuat terhadap faktor f_i daripada korelasi suatu

variabel x_i dengan faktor f_i yang lain. Koefisien matriks korelasi inilah yang disebut sebagai *factor loading* dapat dilihat pada Tabel 4.5.

Tabel 4.5 Matriks Komponen

Variabel	Komponen	
	1	2
Pajak Air Tanah	0,565	0,589
Pajak Hiburan	0,937	
Pajak Penerangan Jalan		0,865
Pajak Hotel	0,841	
Pajak Restoran	0,953	
Pajak Parkir	0,870	
Pajak Bumi dan Bangunan	0,574	0,344

Berdasarkan Tabel 4.5, ada variabel yang berkorelasi sangat kuat terhadap lebih dari satu faktor, yaitu variabel pajak air tanah yang memiliki nilai korelasi sebesar 0,565 terhadap faktor ke-1 dan sebesar 0,589 terhadap faktor ke-2, sehingga sulit untuk menginterpretasikan faktor-faktor tersebut. Oleh karena itu, perlu dilakukan rotasi terhadap matriks komponen. Proses rotasi yang digunakan adalah metode *varimax rotation*. Nilai *loading factor* setelah dilakukan rotasi terhadap matriks komponen dapat dilihat pada Tabel 4.6.

Tabel 4.6 Matriks Komponen Sesudah Rotasi

Variabel	Komponen	
	1	2
Pajak Air Tanah		0,708
Pajak Hiburan	0,956*	
Pajak Penerangan Jalan		0,858*
Pajak Hotel	0,797	
Pajak Restoran	0,922	
Pajak Parkir	0,871	
Pajak Bumi dan Bangunan	0,640	

Tabel 4.6 memperlihatkan distribusi variabel yang lebih jelas dan nyata sehingga faktor-faktor yang terbentuk lebih mudah untuk diinterpretasikan. Ada 5 variabel yang paling dominan di dalam komponen faktor ke-1, yaitu variabel pajak hiburan diikuti oleh pajak restoran, pajak parkir, pajak hotel serta pajak bumi dan bangunan. Variabel yang paling dominan di dalam faktor ke-2 adalah pajak penerangan jalan selain adanya variabel pajak air

tanah. Dengan demikian, variabel penelitian yang semula berjumlah tujuh, telah direduksi menjadi dua faktor komponen yang dapat menjelaskan variabel awal sebesar 73,923%. Berdasarkan keterkaitan variabel yang paling dominan pada masing-masing faktor, untuk selanjutnya faktor ke-1 disebut faktor kebutuhan dasar hidup (air & listrik) dan faktor ke-2 disebut faktor non-kebutuhan dasar hidup.

Tabel 4.6 juga menunjukkan bobot faktor tertinggi dari masing-masing komponen. Variabel yang memiliki bobot faktor tertinggi pada komponen faktor ke-1 adalah variabel pajak hiburan sebesar 95,6% yang mewakili lima variabel yang terdapat pada komponen faktor ke-1. Sedangkan variabel yang memiliki bobot faktor tertinggi pada komponen faktor ke-2 adalah variabel pajak penerangan jalan sebesar 85,8% yang mewakili dua variabel yang terdapat pada komponen faktor ke-2. Variabel yang memiliki bobot tertinggi pada kedua komponen faktor tersebut akan menjadi variabel dalam pengelompokan dua hingga empat kluster kecamatan di Kota Surabaya menggunakan FKM dan FGK.

Dalam analisis kluster FKM ini diperlukan inisiasi jumlah kluster dan fungsi keanggotaan linier naik dengan bantuan program paket analisis data. Jumlah kluster yang harus ditentukan diawal analisis sebanyak 2 hingga 4 kluster dengan fungsi keanggotaan linier naik sebagai derajat keanggotaan yang digunakan sebagai representasi bilangan *fuzzy*.

Tabel 4.7 Hasil Pengelompokan FKM Kota Surabaya

Kluster	Jumlah Kluster		
	2	3	4
1	29	21	29
2	2	2	2
3		8	-
4			-

Hasil pengelompokan analisis kluster FKM yang ditunjukkan pada Tabel 4.7 ini memberikan hasil yang beragam. Pada inisiasi jumlah kluster sebanyak dua dengan bantuan tabel fungsi keanggotaan pada Lampiran 2a, terdapat 29 kecamatan berada

pada kluster ke-1 dan 2 kecamatan pada kluster ke-2. Pada inisiasi jumlah kluster sebanyak tiga dengan bantuan tabel fungsi keanggotaan pada Lampiran 2b, terdapat 21 kecamatan berada pada kluster ke-1, 2 kecamatan pada kluster ke-2 dan 8 kecamatan pada kluster ke-3. Pada inisiasi jumlah kluster sebanyak empat dengan bantuan tabel fungsi keanggotaan pada Lampiran 2c, terdapat 29 kecamatan berada pada kluster ke-1 dan 2 kecamatan pada kluster ke-2.

Dalam analisis kluster FGK ini diperlukan inisiasi jumlah kluster dan fungsi keanggotaan linier naik dengan bantuan program paket analisis data. Jumlah kluster yang harus ditentukan diawal analisis sebanyak 2 hingga 4 kluster dengan fungsi keanggotaan linier naik sebagai derajat keanggotaan yang digunakan sebagai representasi bilangan *fuzzy*.

Tabel 4.8 Hasil Pengelompokan FGK Kota Surabaya

Kluster	Jumlah Kluster		
	2	3	4
1	3	3	7
2	28	2	20
3		26	2
4			2

Hasil pengelompokan analisis kluster FGK yang ditunjukkan pada Tabel 4.8 ini memberikan hasil yang beragam. Pada inisiasi jumlah kluster sebanyak dua dengan bantuan tabel fungsi keanggotaan pada Lampiran 2a, terdapat 3 kecamatan berada pada kluster ke-1 dan 28 kecamatan pada kluster ke-2. Pada inisiasi jumlah kluster sebanyak tiga dengan bantuan tabel fungsi keanggotaan pada Lampiran 2b, terdapat 3 kecamatan berada pada kluster ke-1, 2 kecamatan pada kluster ke-2 dan 26 kecamatan pada kluster ke-3. Pada inisiasi jumlah kluster sebanyak empat dengan bantuan tabel fungsi keanggotaan pada Lampiran 2c, terdapat 7 kecamatan berada pada kluster ke-1, 20 kecamatan pada kluster ke-2, 2 kecamatan pada kluster ke-3 dan 2 kecamatan pada kluster ke-4.

4.5 Evaluasi Kinerja Metode Pengelompokan

Kinerja metode pengelompokan dievaluasi berdasarkan tingkat optimum dan kebaikan hasil *clustering*. Kluster optimum ditentukan dengan bantuan *clusterSim package* berdasarkan kriteria nilai *Pseudo F* tertinggi. Dari hasil pengelompokan kedua metode *clustering* tersebut maka nilai *Pseudo F* ditunjukkan pada Tabel 4.9.

Tabel 4.9 Perhitungan *Pseudo F* Kota Surabaya

Klaster	Metode Pengelompokan	
	Fuzzy K-Means	Fuzzy Gustafson-Kessel
2	836,1309	75,5531
3	473,5886	10,8015
4	836,1309	369,3943

Tabel 4.9 menunjukkan bahwa terdapat perbedaan kelompok yang optimum antara kedua metode *clustering* dengan menggunakan fungsi keanggotaan linier naik. Metode FKM menghasilkan kelompok optimum pada saat penggunaan pengelompokan 2 klaster dengan nilai *Pseudo F* sebesar 836,1. Metode FGK menghasilkan kelompok optimum pada saat penggunaan pengelompokan 4 klaster dengan nilai *Pseudo F* sebesar 369,3. Selanjutnya metode pengelompokan terbaik ditentukan berdasarkan kriteria mencari nilai terkecil pada nilai *icd-rate*. Didapatkan *icd-rate* FKM sebesar 0,0321 sedangkan *icd-rate* FGK sebesar 0,0225. Nilai *icd-rate* yang didapatkan dari FGK *clustering* dengan kelompok optimum sebanyak 4 klaster lebih rendah daripada nilai *icd-rate* yang didapatkan dari FKM *clustering* dengan kelompok optimum sebanyak 2 klaster. Berdasarkan kriteria *icd-rate*, metode FGK *clustering* lebih baik daripada FKM *clustering*.

Tabel 4.10 Profil Pengelompokan Kecamatan di Kota Surabaya

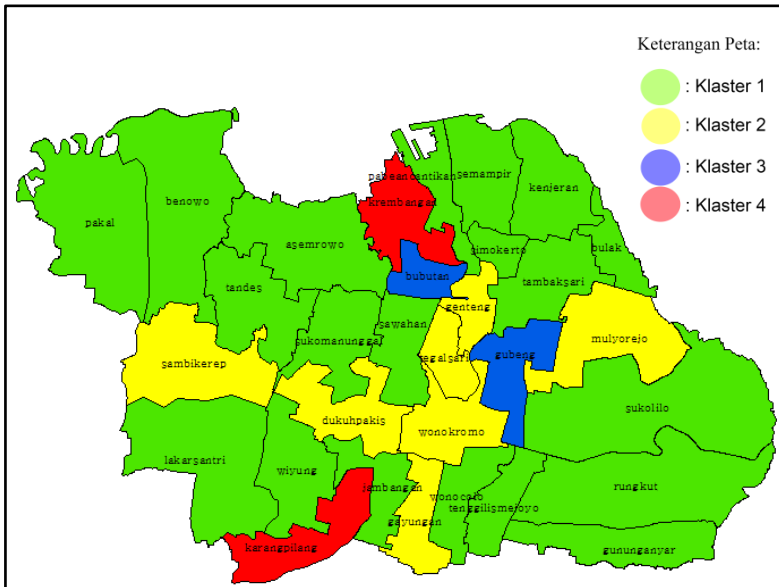
Klaster	Kecamatan
1	Dukuh Pakis, Gayungan, Genteng, Mulyorejo, Sambikerep, Tegalsari, Wonokromo
2	Asemrowo, Benowo, Bulak, Gunung Anyar, Jambangan, Kenjeran, Lakarsantri, Pabean Cantian, Pakal, Rungkut, Sawahan, Semampir, Simokerto, Sukolilo, Sukomanunggal, Tambaksari, Tandes, Tenggilis Mejoyo, Wiyung, Wonocolo
3	Bubutan, Gubeng
4	Karang Pilang, Krembangan

Tabel 4.10 menunjukkan anggota dari setiap klaster yang terbentuk dari pengelompokan dengan metode *fuzzy FGK clustering* dengan fungsi keanggotaan linier naik dengan jumlah klaster sebanyak 4 klaster. Tabel 4.10 menyatakan jumlah kecamatan yang termasuk dalam klaster 1 sebanyak 7 kecamatan, jumlah kecamatan yang termasuk dalam klaster 2 sebanyak 20 kecamatan, jumlah kecamatan yang termasuk dalam klaster 3 sebanyak 2 kecamatan dan jumlah kecamatan yang termasuk dalam klaster 4 sebanyak 2 kecamatan.

4.6 Karakteristik Klaster di Kota Surabaya

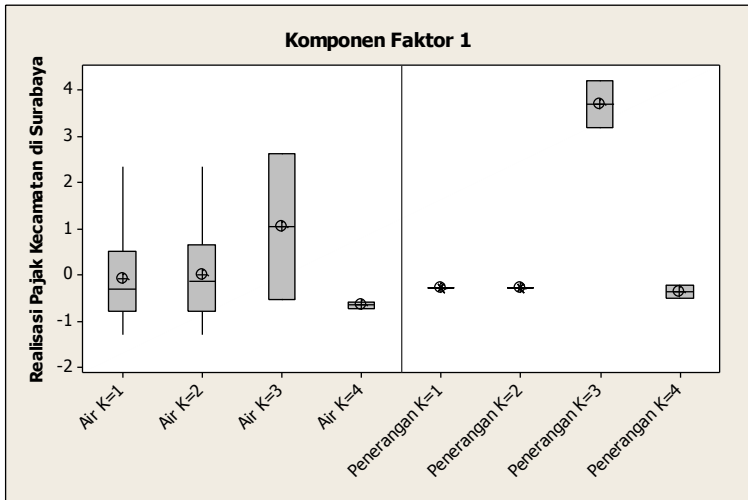
Hasil pengelompokan kecamatan yang didapatkan dari *fuzzy k-means clustering* dan *fuzzy gustafson-kessel clustering* memiliki karakteristik klaster yang beragam. Karakteristik setiap klaster yang terbentuk digunakan untuk membedakan antara suatu klaster dengan klaster yang lain. Diharapkan adanya perbedaan antar klaster tersebut akan dapat menciptakan klaster yang paling unggul jika dibandingkan dengan klaster yang lain. Hasil pengelompokan FGK empat klaster kecamatan di Kota Surabaya telah dinyatakan optimum dan terbaik.

Gambar 4.3 menunjukkan klaster ke-2 adalah klaster dengan warna hijau yang terdiri dari Asemrowo, Benowo, Bulak, Gunung Anyar, Jambangan, Kenjeran, Lakarsantri, Pabean Cantian, Pakal, Rungkut, Sawahan, Semampir, Simokerto, Sukolilo, Sukomanunggal, Tambaksari, Tandes, Tenggilis Mejoyo, Wiyung, Wonocolo. Anggota dari klaster ke-1 adalah Dukuh Pakis, Gayungan, Genteng, Mulyorejo, Sambikerep, Tegalsari, Wonokromo dengan warna kuning. Karang Pilang dan Krembangan merupakan anggota dari klaster ke-3 dengan warna biru. Sedangkan anggota klaster ke-4 dijelaskan dengan warna merah yaitu Bubutan dan Gubeng.



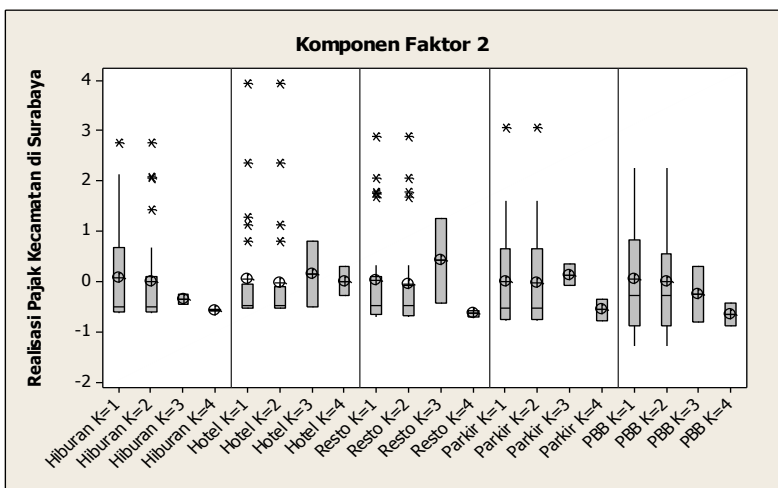
Gambar 4.3. Peta Pengelompokan Kecamatan di Surabaya

Kemudian untuk mengetahui perbedaan karakteristik setiap klaster yang telah terbentuk secara visual pada komponen faktor pertama dapat dilihat pada *boxplot*. Terdapat perbedaan karakteristik antar klaster yang terbentuk. Gambar 4.4 menunjukkan visualisasi dari komponen faktor ke-1 yang terdiri dari pajak air tanah dan pajak penerangan jalan sedangkan. Pajak penerangan jalan dan pajak air tanah yang menjadi anggota klaster ke-3 konsisten mengungguli klaster lainnya pada komponen faktor ke-1.



Gambar 4.4. *Boxplot* Komponen Faktor 1 Kota Surabaya

Gambar 4.5 menunjukkan visualisasi dari komponen faktor ke-2 yang terdiri dari pajak hiburan, hotel, restoran, parkir serta pajak bumi dan bangunan. *Boxplot* yang dihasilkan beragam.



Gambar 4.5. *Boxplot* Komponen Faktor 2 Kota Surabaya

Pajak hiburan, pajak parkir dan PBB yang menjadi anggota klaster ke-1 konsisten mengungguli klaster lainnya pada komponen faktor ke-2. Pajak hotel dan Pajak Restoran yang menjadi anggota klaster ke-3 konsisten mengungguli klaster lainnya pada komponen faktor ke-2. Kemudian karakteristik klaster dibandingkan dengan klaster lain untuk mengetahui gambaran masing-masing variabel pajak daerah menjelaskan setiap klaster yang terbentuk. Berikut analisis statistika deskriptif setiap klaster terbentuk.

Tabel 4.11 Perbandingan Karakteristik Klaster Surabaya (dalam miliar rupiah)

Variabel	Klaster			
	1	2	3	4
X ₁	0,057	0,042	0,074	0,000
X ₂	7,142	0,393	0,837	1,777
X ₃	0,027	0,011	165,995	12,072
X ₄	20,986	2,127	8,777	1,906
X ₅	32,349	3,185	16,703	1,104
X ₆	6,323	0,770	2,469	0,598
X ₇	43,189	23,425	23,191	16,137

Tabel 4.11 menunjukkan nilai rata-rata dari setiap klaster yang terbentuk dari pengelompokan *fuzzy gustafson-kessel clustering* sebanyak empat klaster dengan fungsi keanggotaan linier naik. Klaster ke-2 dan klaster ke-4 tidak memiliki nilai rata-rata tertinggi. Klaster ke-1 memiliki nilai rata-rata tertinggi pada variabel pajak hiburan sebesar 7,142 miliar rupiah, variabel pajak hotel sebesar 20,986 miliar rupiah, variabel pajak restoran sebesar 32,349 miliar rupiah, variabel pajak parkir sebesar 6,323 miliar rupiah dan variabel pajak bumi bangunan sebesar 43,189 miliar rupiah. Klaster ke-3 memiliki nilai rata-rata tertinggi pada variabel pajak air tanah sebesar 74,936 juta rupiah dan pajak penerangan jalan sebesar 165,995 miliar rupiah. Sehingga dapat ditarik keputusan akhir bahwa klaster kecamatan yang paling unggul dari ketujuh variabel pajak daerah adalah klaster ke-1 yang beranggotakan kecamatan Dukuh Pakis, Gayungan, Genteng, Mulyorejo, Sambikerep, Tegalsari, Wonokromo.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis yang telah dilakukan maka dapat ditarik beberapa kesimpulan sebagai berikut:

1. Studi simulasi pengelompokan pada penelitian ini bertujuan untuk menunjukkan tingkat keberhasilan suatu metode dalam melakukan pengelompokan terhadap populasi. Metode FGK secara umum memiliki akurasi yang lebih baik dibandingkan metode FKM dalam melakukan pengelompokan pada seluruh data bangkitan berdistribusi normal multivariat dan normal univariat.
2. Statistika deskriptif yang dibentuk dari tujuh variabel pajak daerah di Kota Surabaya menggambarkan keberagaman tingkat realisasi pajak daerah per kecamatan di Kota Surabaya. Terdapat beberapa kecamatan yang tidak memiliki realisasi pajak yaitu pajak hiburan pada kecamatan Benowo, Kenjeran dan Semampir; pajak penerangan jalan pada kecamatan Benowo, Bulak, Kenjeran dan Pakal; pajak hotel pada kecamatan Benowo, Gunung Anyar, Kenjeran dan Pakal. Sedangkan tingkat realisasi pajak daerah tertinggi dicapai variabel pajak penerangan jalan di kecamatan Gubeng.
3. Analisis Faktor mereduksi 7 variabel metrik menjadi 2 komponen utama, yaitu faktor ke-2 yang disebut faktor non-kebutuhan dasar hidup dengan variabel yang berkaitan yaitu pajak hiburan, pajak hotel, pajak restoran, pajak parkir, pajak bumi dan bangunan kemudian faktor ke-1 disebut faktor kebutuhan dasar hidup, dengan variabel yang berkaitan adalah pajak air tanah dan pajak penerangan jalan.
4. Berdasarkan *Pseudo F-Statistics* didapatkan pengelompokan optimum menggunakan analisis kluster FKM dengan jumlah kluster sebanyak 2 dan analisis kluster FGK dengan jumlah kluster sebanyak 4. Sedangkan metode terbaik berdasarkan *icdrate* terkecil didapatkan oleh analisis kluster FGK dengan

jumlah klaster sebanyak 4. Klaster ke-1 memiliki anggota sebanyak 7 kecamatan yaitu Dukuh Pakis, Gayungan, Genteng, Mulyorejo, Sambikerep, Tegalsari, Wonokromo; klaster ke-2 beranggotakan 20 kecamatan yaitu Asemrowo, Benowo, Bulak, Gunung Anyar, Jambangan, Kenjeran, Lakarsantri, Pabean Cantian, Pakal, Rungkut, Sawahan, Semampir, Simokerto, Sukolilo, Sukomanunggal, Tambaksari, Tandes, Tenggilis Mejoyo, Wiyung, Wonocolo; klaster ke-3 beranggotakan 2 kecamatan yaitu Karang Pilang dan Krembangan dan klaster ke-4 beranggotakan 2 kecamatan yaitu Bubutan dan Gubeng.

5. Karakteristik klaster dijelaskan melalui perbandingan rata-rata klaster di Kota Surabaya dan *Boxplot* dengan klaster ke-1 yang merupakan klaster yang paling unggul dimana setiap anggota pada klaster memiliki rata-rata realisasi pajak yang lebih tinggi daripada realisasi pajak pada anggota klaster 2,3 dan 4.

5.2 Saran

Saran yang dapat disampaikan berdasarkan penelitian ini untuk penelitian selanjutnya adalah sebagai berikut:

1. Penelitian mengenai seberapa besar toleransi jumlah antar dataset yang mampu membedakan akurasi metode FGK dan FKM dapat dieksplorasi oleh penelitian selanjutnya.
2. Berdasarkan observasi lapangan di BPKPD Kota Surabaya, ditemukan variabel lain dengan proporsi jumlah realisasi pajak yang besar terhadap total realisasi pajak daerah Kota Surabaya yaitu proporsi variabel BPHTB (Bea Perolehan Hak atas Tanah dan Bangunan) sehingga penelitian selanjutnya dapat menggunakan variabel BPHTB.
3. Penelitian mengenai klaster unggul yang beranggotakan 7 kecamatan dapat dieksplorasi agar optimalisasi realisasi pajak dapat diwujudkan serta alasan keunggulan klaster tersebut dapat dijelaskan lebih rinci.

DAFTAR PUSTAKA

- Adika, N. (2008). *Pengembangan Wilayah Kabupaten Sidoarjo Sebagai Wilayah Kabupaten Sidoarjo Sebagai Wilayah Pinggiran Kota Metropolitan Surabaya dan Mobilitas Penduduk*.
<http://ojs.unud.ac.id/index.php/piramida/article/view/2974/2132>. Diakses pada tanggal 19 Agustus 2017.
- Alvionita. (2017). *Metode Ensemble ROCK dan SWFM untuk pengelompokan Data Campuran Numerik dan Kategorik pada kasus aksesi jeruk*. Surabaya: Tesis: Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Amalia, N.A., Widodo, D.A., & Oktaviana, P.P. (2016). Analisis Clustering Perusahaan Sub Sektor Perbankan Berdasarkan Rasio Keuangan CAMELS Tahun 2014 Menggunakan Metode Fuzzy C-Means dan Fuzzy Gustafson-Kessel. *Jurnal Sains dan Seni ITS*. 5 (2).
- Andesberg, M. R. (1973). *Cluster Analysis for Applications*. New York: Academic Press.
- Badan Kebijakan Fiskal. (2014). *Ringkasan Eksekutif Kajian Potensi Penerimaan Perpajakan Berdasarkan Pendekatan Makro*. Jakarta: Badan Kebijakan Fiskal Kementerian Keuangan.
- Badan Pendidikan dan Pelatihan Keuangan. (2015). *Pengelolaan Sumber Penerimaan Pajak*.
<http://www.bppk.kemenkeu.go.id/publikasi/artikel/147>.
Diakses pada tanggal 16 Agustus 2017.
- Badan Pusat Statistik, (2015). *Kota Surabaya Dalam Angka 2015*. Surabaya: Badan Pusat Statistik Kota Surabaya.
- Balasko, B., Abonyi, J. & Feil B. (2007). *Fuzzy Clustering and Data Analysis Toolbox For Use with Matlab*. University of Vezprem: Vezprem.
- Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum Press.

- Brotodihardjo, R. S. (1993). *Pengantar Ilmu Hukum Pajak*. Bandung: PT. Eresco.
- Bunkers, M. J. & James, R. M. (1996), Definition of Climate Regions in The Northern Plains Using an Objective Cluster Modification Technique. *J. Climate*. 130-146.
- Dirjen Anggaran Kementerian Keuangan. (2011). *Nota Keuangan dan RAPBN 2011*.
http://www.anggaran.depkeu.go.id/content/APBN%202011_BabV.pdf. Diakses pada 12 Agustus 2017.
- Dirjen Perimbangan Kementerian Keuangan. (2017). *Realisasi Anggaran Pendapatan dan Belanja Daerah 2016*.
<http://www.djpk.depkeu.go.id/?p=4666>. Diakses pada tanggal 9 Oktober 2017.
- Gustafson, D. & Kessel, W.C. (1979). *Fuzzy Clustering with a Fuzzy Covariance Matrix*. San Diego: Paper in Proceedings of the IEEE on Decision and Control. 761 - 766.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. (2009). *Multivariate Data Analysis, 7th edition*. New Jersey: Prentice Hall, Inc.
- Halim, A. & Kusufi, M. S. (2014). *Akuntansi Keuangan Daerah edisi keempat*. Jakarta: Salemba Empat.
- Hartati, S., Hamzah, A. (2005). Kajian Eksperimen Kinerja Fuzzy Clustering C Mean, Guste-Kessel, Gath-Geva dan C Regresi. *Jurnal Pakar Teknologi Informasi dan Bisnis*. 6 (1).
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis, sixth edition*. New Jersey: Pearson Education Inc.
- Kusuma, A. B. (1992). *Risalah Sidang Badan Penyelidik Usaha-Usaha Persiapan Kemerdekaan Indonesia (BPUPKI): Panitia Persiapan Kemerdekaan Indonesia (PPKI): 29 Mei 1945-19 Agustus 1945*. Jakarta: Sekretariat Negara.
- Kusuma, H. B. (2017). *Penentuan Lokasi Baru Kantor Otoritas Jasa Keuangan (OJK) di Pulau Sumatera Menggunakan Analisis Faktor dan Analisis Kluster Fuzzy Gustafson-*

- Kessel. Tugas Akhir: Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Kusumadewi, S., & Hari, P. (2004). *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan*. Yogyakarta: Graha Ilmu.
- Laraswati, T. F. (2014). *Perbandingan Kinerja Metode Complete Linkage, Metode Average Linkage, dan Metode K-Means Dalam Menentukan Hasil Analisis Cluster*. Tugas Akhir: Jurusan Pendidikan Matematika, Universitas Negeri Yogyakarta.
- Law, A.M. and Kelton, W.D. (1991) *Simulation Modelling and Analysis*. 2nd edition. New York: McGraw-Hill.
- Lutfi, R. A. (2016). *Analisa Perubahan Kelompok Kecamatan Berdasarkan Pajak Daerah Tahun 2014-2015 Di Kabupaten Sidoarjo*. Tugas Akhir: Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Mauliyadi, A., Sofyan, H., Subianto, M., (2003). Perbandingan Metode Fuzzy C Means (FCM) dan Fuzzy Gustafson Kessel (FGK) Menggunakan Data Citra Satelit Quickbird. *Jurnal Transenden*.
- Mingoti, S. A. & Lima, J.O. (2006). Comparing SOM Neural Network with Fuzzy C-Means (FCM), C-Means and Traditional Hierarchical Clustering Algorithms. *European Journal of Operational Research*. 174:1742-1759.
- Nurcholish, H. (2005). *Teori dan Praktek Pemerintahan dan Otonomi Daerah*. Jakarta: Gramedia Widiasarana.
- Orphin, A. R. & Kostylev, V. E. (2006). Towards a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats. *Journal Marine Geology*. 225:209-222.
- Rencher, A. (1998). *Multivariate Statistical Inference and Application*, 2nd edition. New York: John Wiley & Sons Inc.
- Rencher, A. (2002). *Methods of Multivariate Analysis*. Canada: John Wiley & Sons, Inc.

- Ruswandi, B. (2008). *Diktat Perkuliahan Praktikum Statistika Multivariat*. Jakarta: Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah Jakarta.
- Sharma, S. (1996). *Applied Multivariate Techniques*. New York: John Wiley & Sons Inc.
- Simamora, B. (2005). *Analisis Multivariat Pemasaran*. Jakarta: Gramedia Pustaka Utama.
- Sirait, R. (2015). *Pengelompokan Kecamatan dan Faktor-Faktor Yang Mempengaruhi Pajak Daerah Kota Surabaya (Studi Kasus Tahun 2006-2014)*. Tugas Akhir: Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Surabaya, B.P.P.K. (2006). Rencana Strategis Badan Perencanaan Pembangunan Kota Surabaya (Renstra-Bappeko) Tahun 2006-2010. *Government of Surabaya*.
- Republik Indonesia, (2009). Undang-Undang No 28 Tahun 2009 tentang Pajak. Jakarta: Sekretariat Negara.
- Velichkov, N., & Stefanova, K. (2017). *Tax Models in the EU: a Cluster Analysis*. *Economic Alternatives*, (4), 573-583.
- Walpole, R.E. (1998). *Pengantar Metode Statistika* (3rd ed). (B. Sumantri, Penerj.) Jakarta: PT. Gramedia Pustaka Utama.
- Walpole, R. E., Raymond H. Myers, Shaton L. Myers, & Keying Ye. (2012). *Probability & Statistics for Engineers & Scientist 9th Edition*. Boston: Pearson Education, Inc.
- Widodo, A. (2012). *Perbandingan Metode Fuzzy C-Means Clustering dan Fuzzy C-Shell Clustering (Studi Kasus: Kabupaten/Kota di Pulau Jawa Berdasarkan Variabel Pembentuk Indeks Pembangunan Manusia)*. Tesis: Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Zadeh, L. A. (1965). Information and control. *Fuzzy sets*, 8(3), 338-353.

LAMPIRAN

Lampiran 1. Data Penelitian

a. Data Realisasi Pajak Daerah Kota Surabaya

Kecamatan	X ₁	X ₂	...	X ₇
Asemrowo	34442466	1250677	...	35724586536
Benowo	15170336	0	...	27841864537
Bubutan	33154884	527547330	...	13708966871
Bulak	21285716	815828855	...	5818957450
Dukuh Pakis	33808469	8496077250	...	60775283804
Gayungan	33444014	3986945547	...	18691804920
Genteng	109504084	10481584241	...	41213045676
Gubeng	116717127	1147903482	...	32673422651
Gunung Anyar	91336590	71161618	...	12756481607
Jambangan	12905676	61853415	...	5332563843
Karang Pilang	31815857	213401441	...	12109609734
Kenjeran	48108163	0	...	7455914685
Kremlangan	28092350	142050089	...	20165052452
Lakarsantri	22321391	269405363	...	53221373254
Mulyorejo	59207521	6339056067	...	66432706169
Pabean Cantian	75233492	692310706	...	22430822032
Pakal	27521385	143687344	...	9059001453
Rungkut	52478026	79627572	...	33327454859
Sambikerep	21843316	4049724277	...	55506432271
Sawahan	45865782	1291861411	...	18556887675
Semampir	41583977	0	...	10974009261
Simokerto	29235491	544408247	...	7422505652
Sukolilo	59930008	154437046	...	55026986941
Sukomanunggal	36593139	244176415	...	43937293211
Tambaksari	62356871	1672402177	...	22412009054
Tandes	70800252	58236188	...	19206614680
Tegalsari	73306323	8383220053	...	34755809411
Tenggiling Mejoyo	48190175	108782119	...	20533135887
Wiyung	23002580	712343540	...	34294315827
Wonocolo	34212040	939205969	...	23181557023
Wonokromo	72141899	8260793144	...	24949457693

b. Sampel Data Bangkitan Normal Multivariat 30 pengamatan

No	X_1	X_2	No	X_1	X_2
1	8,269	8,830	1	8,318	8,169
2	9,257	8,752	2	7,669	7,761
3	10,065	10,762	3	10,845	11,257
4	9,575	9,892	4	9,948	8,791
5	9,733	10,151	5	8,689	8,875
6	9,304	9,609	6	10,708	10,317
7	8,537	8,993	7	8,695	8,974
8	10,060	9,953	8	10,221	10,283
9	8,001	8,751	9	7,933	8,114
10	7,626	7,809	10	6,550	6,699
11	9,551	9,154	11	8,824	9,098
12	9,859	9,547	12	9,308	8,634
13	11,306	11,539	13	7,925	7,720
14	7,781	8,299	14	10,192	10,664
15	8,695	8,645	15	10,096	10,008
16	9,413	9,069	16	9,078	8,701
17	8,503	8,564	17	9,278	9,286
18	8,201	8,550	18	8,916	8,519
19	7,956	7,417	19	9,770	8,966
20	7,094	7,906	20	7,371	7,607
21	7,868	7,770	21	9,063	8,916
22	8,154	9,098	22	9,967	9,852
23	9,494	9,463	23	10,089	8,911
24	8,383	8,074	24	9,063	7,923
25	10,315	10,005	25	7,399	7,325
26	7,624	7,302	26	6,329	6,633
27	7,243	6,999	27	7,881	8,368
28	7,569	7,744	28	10,450	10,558
29	8,487	8,444	29	8,865	7,883
30	9,073	8,905	30	11,172	11,306

c. Sampel Data Bangkitan Normal Multivariat 60 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	89,203	90,141	1	88,593	89,311
2	89,018	89,165	2	91,200	90,377
3	88,566	90,086	3	89,251	89,683
4	91,297	91,405	4	90,109	90,806
5	92,325	91,102	5	88,916	89,683
:	:	:	:	:	:
56	89,920	90,040	56	89,486	88,743
57	91,377	90,954	57	89,483	89,701
58	88,504	86,969	58	90,202	88,374
59	90,490	88,877	59	88,340	88,050
60	90,356	89,153	60	89,868	88,408

d. Sampel Data Bangkitan Normal Multivariat 90 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	90,085	91,010	1	89,524	89,045
2	88,823	89,236	2	88,690	88,879
3	89,723	89,178	3	92,197	91,682
4	90,423	90,950	4	90,959	89,740
5	87,699	88,412	5	89,503	90,084
:	:	:	:	:	:
86	88,912	88,305	86	89,861	89,665
87	90,362	88,644	87	90,106	88,452
88	88,746	86,804	88	89,649	88,654
89	90,583	89,162	89	89,877	89,969
90	91,851	90,884	90	89,341	88,634

e. Sampel Data Bangkitan Normal Multivariat 120 Pengamatan

	V1	V2		V1	V2
1	89,897	89,584	1	90,235	91,924
2	88,842	89,506	2	91,081	90,551
3	88,801	88,114	3	88,494	89,075
4	89,250	89,496	4	88,108	88,862
5	91,054	89,398	5	91,466	90,487
:	:	:	:	:	:
116	88,791	87,534	116	89,119	88,522
117	90,575	90,102	117	90,191	90,591
118	89,835	87,933	118	91,648	89,167
119	88,657	89,109	119	90,095	89,250
120	89,681	88,884	120	90,145	89,359

f. Sampel Data Bangkitan Normal Multivariat 150 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	89,891	88,683	1	90,244	88,193
2	89,142	89,981	2	91,196	90,025
3	89,864	88,051	3	90,239	90,734
4	89,525	89,685	4	89,534	90,166
5	88,930	89,522	5	89,904	89,395
:	:	:	:	:	:
146	89,808	88,062	146	90,522	89,846
147	90,751	90,622	147	87,409	86,639
148	89,570	88,403	148	90,974	90,058
149	90,432	88,526	149	91,107	90,138
150	91,390	90,053	150	90,814	90,954

g. Sampel Data Bangkitan Normal Multivariat 300 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	88,383	88,062	1	90,070	89,843
2	90,101	89,080	2	90,652	91,867
3	90,581	90,147	3	90,479	89,640
4	89,215	89,974	4	90,282	91,011
5	91,338	89,872	5	88,083	88,810
:	:	:	:	:	:
296	88,549	87,744	296	88,651	88,442
297	90,185	88,247	297	90,245	89,446
298	89,644	89,402	298	89,068	89,032
299	90,967	90,123	299	90,738	88,327
300	90,802	89,761	300	87,859	88,919

h. Sampel Data Bangkitan Normal Multivariat 1500 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	90,554	90,395	1	89,919	89,957
2	90,264	90,222	2	89,881	90,083
3	90,856	91,201	3	87,734	89,345
4	90,263	89,693	4	89,211	90,620
5	89,635	90,222	5	88,309	88,790
:	:	:	:	:	:
1496	89,251	89,217	1496	88,921	88,136
1497	90,899	89,252	1497	87,422	87,976
1498	89,899	88,269	1498	89,325	89,099
1499	91,183	90,356	1499	89,745	89,585
1500	90,682	89,235	1500	90,014	88,864

i. Sampel Data Bangkitan Normal Univariat 30 pengamatan

No	X_1	X_2	No	X_1	X_2
1	6,609	7,390	1	8,134	13,435
2	4,003	7,604	2	8,804	7,766
3	9,581	9,758	3	11,774	11,264
4	6,361	9,930	4	6,379	8,020
5	4,596	8,040	5	7,642	8,470
6	8,328	11,947	6	11,400	9,257
7	13,873	7,926	7	10,997	14,540
8	7,680	8,373	8	7,884	8,842
9	9,927	7,600	9	9,382	7,849
10	8,876	12,267	10	7,127	10,755
11	9,287	7,927	11	12,436	10,525
12	7,346	8,114	12	10,585	5,626
13	9,519	8,944	13	8,640	5,982
14	4,392	10,347	14	6,457	9,634
15	7,232	9,123	15	9,724	9,910
16	12,242	9,212	16	7,833	5,602
17	7,117	11,235	17	9,168	4,853
18	10,999	9,736	18	10,264	9,589
19	12,884	6,533	19	4,716	9,587
20	5,934	9,923	20	6,185	10,708
21	7,145	11,882	21	7,601	10,414
22	9,919	8,321	22	8,227	7,268
23	8,264	6,785	23	6,512	10,686
24	12,003	7,857	24	9,873	6,779
25	6,151	10,254	25	6,836	9,471
26	9,839	11,636	26	9,591	9,085
27	10,335	11,207	27	9,728	11,195
28	10,761	9,053	28	8,494	7,283
29	9,012	8,532	29	8,537	9,990
30	7,682	10,616	30	7,750	4,653

j. Sampel Data Bangkitan Normal Univariat 60 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	78,129	97,450	1	90,230	97,593
2	74,744	75,555	2	131,467	76,991
3	87,801	80,539	3	60,057	83,436
4	94,565	128,385	4	95,578	49,345
5	69,676	95,054	5	81,143	93,358
:	:	:	:	:	:
56	79,063	74,390	56	98,782	75,559
57	103,735	103,252	57	96,382	105,694
58	96,376	101,949	58	112,387	76,586
59	109,488	62,503	59	73,328	100,388
60	138,187	79,209	60	89,280	91,571

k. Sampel Data Bangkitan Normal Univariat 90 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	68,624	79,231	1	90,315	73,928
2	88,738	89,605	2	110,349	103,613
3	105,532	98,359	3	66,031	115,051
4	83,651	106,916	4	109,019	61,586
5	70,444	32,202	5	114,451	59,293
:	:	:	:	:	:
86	108,224	86,239	86	95,615	51,746
87	65,954	42,918	87	47,856	76,651
88	72,786	50,143	88	100,227	112,961
89	100,344	94,160	89	82,531	84,156
90	117,535	78,475	90	129,069	95,325

l. Sampel Data Bangkitan Normal Univariat 120 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	99,045	82,469	1	82,213	106,496
2	89,772	94,598	2	121,496	91,861
3	112,987	124,503	3	85,657	88,225
4	80,017	104,526	4	69,978	97,031
5	120,584	90,922	5	82,041	97,405
:	:	:	:	:	:
116	93,192	117,550	116	41,090	75,065
117	59,926	122,387	117	64,059	90,995
118	78,420	104,454	118	91,300	93,817
119	107,120	79,369	119	99,224	83,326
120	79,400	72,624	120	80,168	75,747

m. Sampel Data Bangkitan Normal Univariat 150 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	113,417	115,781	1	85,775	4,787
2	107,697	105,324	2	79,615	78,528
3	63,642	90,644	3	96,143	89,664
4	57,135	75,983	4	102,719	83,910
5	111,185	103,528	5	47,996	117,602
:	:	:	:	:	:
146	31,301	134,910	146	106,465	88,707
147	42,306	80,917	147	102,794	70,668
148	54,382	93,897	148	71,932	97,517
149	105,500	90,630	149	102,083	49,584
150	68,099	54,213	150	117,000	94,896

n. Sampel Data Bangkitan Normal Univariat 300 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	87,648	106,824	1	125,028	91,070
2	91,615	92,827	2	92,157	96,776
3	85,114	71,692	3	83,859	105,469
4	75,216	65,362	4	112,072	100,497
5	102,885	75,033	5	111,071	84,948
:	:	:	:	:	:
296	127,342	90,259	296	67,852	81,122
297	65,228	49,846	297	66,986	116,596
298	73,695	99,432	298	98,752	78,362
299	83,762	82,195	299	50,312	116,125
300	81,330	84,613	300	72,838	44,108

o. Sampel Data Bangkitan Normal Univariat 1500 Pengamatan

No	X_1	X_2	No	X_1	X_2
1	83,638	78,039	1	73,055	119,494
2	83,690	87,912	2	91,409	97,092
3	107,686	69,503	3	121,624	102,082
4	52,292	58,754	4	90,412	93,859
5	104,644	61,084	5	87,533	63,084
:	:	:	:	:	:
1496	71,599	68,209	1496	76,618	79,270
1497	148,051	107,580	1497	75,707	99,187
1498	127,351	122,161	1498	108,060	107,540
1499	76,057	109,510	1499	114,320	79,162
1500	88,809	91,898	1500	104,769	80,951

Lampiran 2. Output KMO dan Bartlett Test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,671
Approx. Chi-Square		147,954
Bartlett's Test of Sphericity		Df
		21
		Sig.
		,000

Lampiran 3. Output Nilai Communalities

Communalities		
	Initial	Extraction
Air Tanah	1,000	,667
Hiburan	1,000	,916
Penerangan Jalan	1,000	,754
Hotel	1,000	,713
Restoran	1,000	,908
Parkir	1,000	,769
PBB	1,000	,448
Extraction Method: Principal Component Analysis.		

Lampiran 4. Output Scree Plot

Lampiran 5. Tabel *Membership Function*

a. Fungsi Keanggotaan 2 Klaster

No	U_1	U_2
1	0,080530421	0,919469579
2	0,05825636	0,94174364
3	0,655974232	0,344025768
4	0,002132481	0,997867519
5	0,31000956	0,68999044
6	0,071862246	0,928137754
7	0,285297319	0,714702681
8	1	0
9	0,029362874	0,970637126
10	0	1
11	0,120034552	0,879965448
12	0,003225509	0,996774491
13	0,051937894	0,948062106
14	0,159492908	0,840507092
15	0,297721162	0,702278838
16	0,068143365	0,931856635
17	0,007604444	0,992395556
18	0,083441612	0,916558388
19	0,18253917	0,81746083
20	0,090220903	0,909779097
21	0,013032686	0,986967314
22	0,010760438	0,989239562
23	0,147382022	0,852617978
24	0,140388213	0,859611787
25	0,05007741	0,94992259
26	0,037605316	0,962394684
27	0,359229925	0,640770075
28	0,048319626	0,951680374
29	0,097237465	0,902762535
30	0,083530723	0,916469277
31	0,200335251	0,799664749

b. Fungsi Keanggotaan 3 Klaster

No	U_1	U_2	U_3
1	0,965111038	0,017444481	0,017444481
2	0,698169159	0,150915421	0,150915421
3	1	0	0
4	0,025556564	0,487221718	0,487221718
5	1	0	0
6	0,861227916	0,069386042	0,069386042
7	1	0	0
8	1	0	0
9	0,351897252	0,324051374	0,324051374
10	0	0,5	0,5
11	0,039924287	0,480037857	0,480037857
12	0,038655886	0,480672057	0,480672057
13	0,622445961	0,18877702	0,18877702
14	1	0	0
15	1	0	0
16	0,816659257	0,091670372	0,091670372
17	0,09113491	0,454432545	0,454432545
18	0	0,5	0,5
19	1	0	0
20	0,007396464	0,496301768	0,496301768
21	0,156189288	0,421905356	0,421905356
22	0,12895769	0,435521155	0,435521155
23	0,069761415	0,465119292	0,465119292
24	0,062130904	0,468934548	0,468934548
25	0,60014912	0,19992544	0,19992544
26	0,450678207	0,274660897	0,274660897
27	1	0	0
28	0,579083086	0,210458457	0,210458457
29	0,0150518	0,4924741	0,4924741
30	0,0000972	0,4999513	0,4999513
31	1	0	0

c. Fungsi Keanggotaan 4 Klaster

No	U_1	U_2	U_3	U_4
1	0,08053042	0,30648986	0,30648986	0,30648986
2	0,05825636	0,31391455	0,31391455	0,31391455
3	0,65597423	0,11467526	0,11467526	0,11467526
4	0,00213248	0,33262251	0,33262251	0,33262251
5	0,31000956	0,22999681	0,22999681	0,22999681
6	0,07186225	0,30937925	0,30937925	0,30937925
7	0,28529732	0,23823423	0,23823423	0,23823423
8	1	0	0	0
9	0,02936287	0,32354571	0,32354571	0,32354571
10	0	0,33333333	0,33333333	0,33333333
11	0,12003455	0,29332182	0,29332182	0,29332182
12	0,00322551	0,33225816	0,33225816	0,33225816
13	0,05193789	0,3160207	0,3160207	0,3160207
14	0,15949291	0,28016903	0,28016903	0,28016903
15	0,29772116	0,23409295	0,23409295	0,23409295
16	0,06814337	0,31061888	0,31061888	0,31061888
17	0,00760444	0,33079852	0,33079852	0,33079852
18	0,08344161	0,30551946	0,30551946	0,30551946
19	0,18253917	0,27248694	0,27248694	0,27248694
20	0,0902209	0,3032597	0,3032597	0,3032597
21	0,01303269	0,32898911	0,32898911	0,32898911
22	0,01076044	0,32974652	0,32974652	0,32974652
23	0,14738202	0,28420599	0,28420599	0,28420599
24	0,14038821	0,28653726	0,28653726	0,28653726
25	0,05007741	0,31664086	0,31664086	0,31664086
26	0,03760532	0,32079823	0,32079823	0,32079823
27	0,35922993	0,21359003	0,21359003	0,21359003
28	0,04831963	0,31722679	0,31722679	0,31722679
29	0,09723747	0,30092085	0,30092085	0,30092085
30	0,08353072	0,30548976	0,30548976	0,30548976
31	0,20033525	0,26655492	0,26655492	0,26655492

d. Fungsi Keanggotaan 30 Pengamatan

No	U_1	U_2
1	0,39459128	0,60540872
2	0,143383989	0,856616011
3	0,332518128	0,667481872
4	0,499378566	0,500621434
5	0,288357427	0,711642573
6	0,952987232	0,047012768
7	0,455522654	0,544477346
8	0,341634407	0,658365593
9	0,064683525	0,935316475
10	0,398664135	0,601335865
11	0,119760676	0,880239324
12	0,480297571	0,519702429
13	0	1
14	0,506015046	0,493984954
15	0,29734251	0,70265749
16	0,597061473	0,402938527
17	0,392985063	0,607014937
18	0,169065597	0,830934403
19	0,529500302	0,470499698
20	0,537104887	0,462895113
21	1	0
22	0,483794651	0,516205349
23	0,446943071	0,553056929
24	0,324020291	0,675979709
25	0,482118349	0,517881651
26	0,03503047	0,96496953
27	0,499072208	0,500927792
28	0,291759065	0,708240935
29	0,459522234	0,540477766
30	0,558116854	0,441883146

e. Fungsi Keanggotaan 60 Pengamatan

No	U_1	U_2
1	0,704179539	0,295820461
2	0,518347512	0,481652488
3	0,700481694	0,299518306
4	0,924922237	0,075077763
5	0,855696554	0,144303446
:	:	:
56	0,67702859	0,32297141
57	0,837291553	0,162708447
58	0,10151346	0,89848654
59	0,447281711	0,552718289
60	0,501855708	0,498144292

f. Fungsi Keanggotaan 90 Pengamatan

No	U_1	U_2
1	0,346607386	0,653392614
2	0,319487792	0,680512208
3	0,804275018	0,195724982
4	0,465209489	0,534790511
5	0,529696587	0,470303413
:	:	:
86	0,454950661	0,545049339
87	0,240560403	0,759439597
88	0,277258237	0,722741763
89	0,508383194	0,491616806
90	0,27464449	0,72535551

g. Fungsi Keanggotaan 120 Pengamatan

No	U_1	U_2
1	0,574330113	0,425669887
2	0,571687967	0,428312033
3	0,273186317	0,726813683
4	0,564134195	0,435865805
5	0,518809177	0,481190823
:	:	:
116	0,148824524	0,851175476
117	0,67640074	0,32359926
118	0,220362949	0,779637051
119	0,489019837	0,510980163
120	0,426807218	0,573192782

h. Fungsi Keanggotaan 150 Pengamatan

No	U_1	U_2
1	0,258701625	0,741298375
2	0,513901558	0,486098442
3	0,144061013	0,855938987
4	0,45042307	0,54957693
5	0,435491063	0,564508937
:	:	:
146	0,147472544	0,852527456
147	0,591156109	0,408843891
148	0,21561811	0,78438189
149	0,216636779	0,783363221
150	0,471510655	0,528489345

i. Fungsi Keanggotaan 300 Pengamatan

No	U_1	U_2
1	0,869229287	0,130770713
2	0,647939013	0,352060987
3	0,720499888	0,279500112
4	0,992730432	0,007269568
5	0,908479656	0,091520344
:	:	:
296	0,60609312	0,39390688
297	0,220209542	0,779790458
298	0,582948127	0,417051873
299	0,613852043	0,386147957
300	0,411966811	0,588033189

j. Fungsi Keanggotaan 1500 Pengamatan

No	U_1	U_2
1	0,335062426	0,664937574
2	0,660770531	0,339229469
3	0,822837559	0,177162441
4	0,242142662	0,757857338
5	0,766039263	0,233960737
:	:	:
1496	0,613059814	0,386940186
1497	0,393093777	0,606906223
1498	0,28229593	0,71770407
1499	0,494416391	0,505583609
1500	0,494416391	0,505583609

Lampiran 6. Simulasi FGK Normal Multivariat

a. Sampel Hasil Normal Multivariat 30 pengamatan

Data ke-	FGK	Aktual
1	1	1
2	2	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	2	1
9	1	1
10	2	1
11	2	1
12	2	1
13	1	1
14	1	1
15	2	1
16	2	2
17	2	2
18	1	2
19	2	2
20	1	2
21	2	2
22	1	2
23	2	2
24	2	2
25	2	2
26	2	2
27	2	2
28	2	2
29	2	2
30	2	2

b. Sampel Hasil Normal Multivariat 60 Pengamatan

Data ke-	FGK	Aktual
1	1	1
2	1	1
3	1	1
4	2	1
5	1	1
:	:	:
56	2	2
57	1	2
58	2	2
59	2	2
60	2	2

c. Sampel Hasil Normal Multivariat 90 Pengamatan

Data ke-	FGK	Aktual
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
:	:	:
86	1	2
87	2	2
88	2	2
89	2	2
90	2	2

d. Sampel Hasil Normal Multivariat 120 Pengamatan

Data ke-	FGK	Aktual
1	1	1
2	1	1
3	2	1
4	1	1
5	2	1
:	:	:
116	2	2
117	1	2
118	2	2
119	1	2
120	2	2

e. Sampel Hasil Normal Multivariat 150 Pengamatan

Data ke-	FGK	Aktual
1	2	1
2	1	1
3	2	1
4	1	1
5	1	1
:	:	:
146	2	2
147	1	2
148	2	2
149	2	2
150	2	2

f. Sampel Hasil Normal Multivariat 300 Pengamatan

Data ke-	FGK	Aktual
1	2	1
2	2	1
3	1	1
4	1	1
5	2	1
:	:	:
296	2	2
297	1	2
298	2	2
299	2	2
300	2	2

g. Sampel Hasil Normal Multivariat 1500 Pengamatan

Data ke-	FGK	Aktual
1	1	1
2	1	1
3	1	1
4	2	1
5	1	1
:	:	:
1496	2	2
1497	2	2
1498	1	2
1499	2	2
1500	2	2

Lampiran 7. Simulasi FGK Normal Univariat

a. Sampel Hasil Normal Univariat 30 pengamatan

Data ke-	FGK	Aktual
1	2	1
2	1	1
3	2	1
4	1	1
5	1	1
6	1	1
7	2	1
8	2	1
9	2	1
10	1	1
11	2	1
12	2	1
13	2	1
14	1	1
15	1	1
16	2	2
17	1	2
18	2	2
19	2	2
20	1	2
21	1	2
22	2	2
23	2	2
24	2	2
25	1	2
26	1	2
27	1	2
28	2	2
29	2	2
30	1	2

b. Sampel Hasil Normal Univariat 60 Pengamatan

Data ke-	FGK	Aktual
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
:	:	:
56	1	2
57	2	2
58	2	2
59	2	2
60	2	2

c. Sampel Hasil Normal Univariat 90 Pengamatan

Data ke-	FGK	Aktual
1	2	1
2	1	1
3	1	1
4	2	1
5	2	1
:	:	:
86	2	2
87	2	2
88	2	2
89	1	2
90	1	2

d. Sampel Hasil Normal Univariat 120 Pengamatan

Data ke-	FGK	Aktual
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
:	:	:
116	2	2
117	1	2
118	1	2
119	1	2
120	2	2

e. Sampel Data Bangkitan Normal Univariat 150 Pengamatan

Data ke-	FGK	Aktual
1	1	1
2	1	1
3	2	1
4	1	1
5	2	1
:	:	:
146	1	2
147	2	2
148	2	2
149	2	2
150	1	2

f. Sampel Hasil Normal Univariat 300 Pengamatan

Data ke-	FGK	Aktual
1	1	1
2	1	1
3	2	1
4	2	1
5	2	1
:	:	:
296	2	2
297	2	2
298	1	2
299	1	2
300	2	2

g. Sampel Hasil Normal Univariat 1500 Pengamatan

Data ke-	FGK	Aktual
1	2	1
2	2	1
3	1	1
4	2	1
5	1	1
:	:	:
1496	2	2
1497	2	2
1498	1	2
1499	1	2
1500	1	2

Lampiran 8. Simulasi FKM Normal Multivariat

a. Sampel Hasil Normal Multivariat 30 pengamatan

Data ke-	FKM	Aktual
1	1	1
2	2	1
3	2	1
4	2	1
5	2	1
6	2	1
7	1	1
8	2	1
9	1	1
10	1	1
11	2	1
12	2	1
13	2	1
14	1	1
15	1	1
16	2	2
17	1	2
18	1	2
19	1	2
20	1	2
21	1	2
22	1	2
23	2	2
24	1	2
25	2	2
26	1	2
27	1	2
28	1	2
29	1	2
30	2	2

b. Sampel Hasil Normal Multivariat 60 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	1	1
3	1	1
4	2	1
5	1	1
:	:	:
56	2	2
57	1	2
58	2	2
59	2	2
60	2	2

c. Sampel Hasil Normal Multivariat 90 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	2	1
3	2	1
4	1	1
5	2	1
:	:	:
86	2	2
87	2	2
88	2	2
89	1	2
90	1	2

d. Sampel Hasil Normal Multivariat 120 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	2	1
3	2	1
4	2	1
5	1	1
:	:	:
116	2	2
117	1	2
118	2	2
119	2	2
120	2	2

e. Sampel Hasil Normal Multivariat 150 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
:	:	:
146	1	2
147	2	2
148	1	2
149	1	2
150	2	2

f. Sampel Hasil Normal Multivariat 300 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	1	1
3	2	1
4	1	1
5	2	1
:	:	:
296	1	2
297	1	2
298	1	2
299	2	2
300	2	2

g. Sampel Hasil Normal Multivariat 1500 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
:	:	:
1496	1	2
1497	2	2
1498	1	2
1499	1	2
1500	1	2

Lampiran 9. Simulasi FKM Normal Univariat

a. Sampel Hasil Normal Univariat 30 pengamatan

Data ke-	FKM	Aktual
1	1	1
2	1	1
3	2	1
4	1	1
5	1	1
6	1	1
7	2	1
8	1	1
9	2	1
10	1	1
11	2	1
12	1	1
13	2	1
14	1	1
15	1	1
16	2	2
17	1	2
18	2	2
19	2	2
20	1	2
21	1	2
22	2	2
23	2	2
24	2	2
25	1	2
26	2	2
27	2	2
28	2	2
29	2	2
30	1	2

b. Sampel Hasil Normal Univariat 60 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	2	1
3	2	1
4	2	1
5	2	1
:		:
56	2	2
57	1	2
58	2	2
59	1	2
60	2	2

c. Sampel Hasil Normal Univariat 90 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	2	1
3	2	1
4	2	1
5	1	1
:		:
86	2	2
87	1	2
88	2	2
89	1	2
90	1	2

d. Sampel Hasil Normal Univariat 120 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	2	1
3	1	1
4	2	1
5	2	1
:	:	:
116	1	2
117	1	2
118	1	2
119	2	2
120	1	2

e. Sampel Data Bangkitan Normal Univariat 150 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	2	1
3	1	1
4	1	1
5	1	1
:	:	:
146	1	2
147	1	2
148	2	2
149	2	2
150	1	2

f. Sampel Hasil Normal Univariat 300 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	1	1
3	1	1
4	1	1
5	2	1
:	:	:
296	2	2
297	2	2
298	1	2
299	2	2
300	1	2

g. Sampel Hasil Normal Univariat 1500 Pengamatan

Data ke-	FKM	Aktual
1	1	1
2	1	1
3	2	1
4	2	1
5	2	1
:	:	:
1496	1	2
1497	1	2
1498	1	2
1499	1	2
1500	2	2

Lampiran 10. Tabel Akurasi Normal Multivariat

output ke-	c=30		c=60		c=90		c=120	
	FGK	FKM	FGK	FKM	FGK	FKM	FGK	FKM
1	0,7	0,333	0,8	0,55	0,711	0,544	0,741	0,55
2	0,7	0,333	0,816	0,55	0,488	0,444	0,7	0,666
3	0,4	0,366	0,816	0,766	0,711	0,544	0,675	0,458
4	0,6	0,466	0,566	0,533	0,744	0,288	0,683	0,458
5	0,7	0,333	0,733	0,533	0,6	0,388	0,741	0,453
6	0,7	0,333	0,633	0,383	0,666	0,566	0,675	0,458
7	0,3	0,333	0,833	0,616	0,677	0,7	0,675	0,458
8	0,7	0,333	0,833	0,616	0,722	0,633	0,683	0,346
9	0,7	0,333	0,65	0,566	0,788	0,366	0,75	0,65
10	0,7	0,333	0,65	0,65	0,788	0,644	0,646	0,391

Lampiran 11. Tabel Akurasi Normal Multivariat c-Data Besar

output ke-	c=150		c=300		c=1500	
	FGK	FKM	FGK	FKM	FGK	FKM
1	0,74	0,466	0,73	0,39	0,778	0,63
2	0,546	0,44	0,79	0,646	0,746	0,625
3	0,733	0,413	0,703	0,61	0,756	0,601
4	0,72	0,493	0,7	0,61	0,709	0,382
5	0,72	0,426	0,7	0,61	0,758	0,617
6	0,72	0,546	0,7	0,61	0,754	0,372
7	0,706	0,373	0,733	0,666	0,748	0,65
8	0,76	0,66	0,773	0,583	0,731	0,596
9	0,726	0,6	0,733	0,593	0,768	0,612
10	0,76	0,586	0,723	0,646	0,754	0,61

Lampiran 12. Tabel Akurasi Normal Univariat

output ke-	c=30		c=60		c=90		c=120	
	FGK	FKM	FGK	FKM	FGK	FKM	FGK	FKM
1	0,7	0,333	0,8	0,55	0,711	0,544	0,741	0,55
2	0,7	0,333	0,816	0,55	0,488	0,444	0,7	0,666
3	0,4	0,366	0,816	0,766	0,711	0,544	0,675	0,458
4	0,6	0,466	0,566	0,533	0,744	0,288	0,683	0,458
5	0,7	0,333	0,733	0,533	0,6	0,388	0,741	0,453
6	0,7	0,333	0,633	0,383	0,666	0,566	0,675	0,458
7	0,3	0,333	0,833	0,616	0,677	0,7	0,675	0,458
8	0,7	0,333	0,833	0,616	0,722	0,633	0,683	0,346
9	0,7	0,333	0,65	0,566	0,788	0,366	0,75	0,65
10	0,7	0,333	0,65	0,65	0,788	0,644	0,646	0,391

Lampiran 13. Tabel Akurasi Normal Univariat c-Data Besar

output ke-	c=150		c=300		c=1500	
	FGK	FKM	FGK	FKM	FGK	FKM
1	0,493	0,5	0,45	0,56	0,534	0,508
2	0,533	0,513	0,5	0,496	0,51	0,49
3	0,54	0,446	0,506	0,503	0,514	0,474
4	0,52	0,46	0,51	0,483	0,515	0,49
5	0,54	0,473	0,51	0,493	0,507	0,504
6	0,473	0,473	0,503	0,466	0,508	0,498
7	0,466	0,54	0,503	0,466	0,516	0,486
8	0,533	0,513	0,56	0,523	0,49	0,495
9	0,533	0,622	0,493	0,496	0,49	0,501
10	0,566	0,622	0,49	0,516	0,516	0,505

Lampiran 14. Hasil Pengelompokan FKM Kota Surabaya

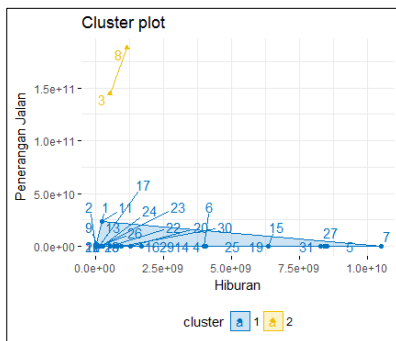
Kecamatan	2 Klaster	3 Klaster	4 Klaster
Asemrowo	1	1	1
Benowo	1	1	1
Bubutan	2	2	2
Bulak	1	1	1
Dukuh Pakis	1	3	1
Gayungan	1	3	1
Genteng	1	3	1
Gubeng	2	2	2
Gunung Anyar	1	1	1
Jambangan	1	1	1
Karang Pilang	1	3	1
Kenjeran	1	1	1
Krembangan	1	1	1
Lakarsantri	1	1	1
Mulyorejo	1	3	1
Pabean Cantian	1	1	1
Pakal	1	1	1
Rungkut	1	1	1
Sambikerep	1	3	1
Sawahan	1	1	1
Semampir	1	1	1
Simokerto	1	1	1
Sukolilo	1	1	1
Sukomanunggal	1	1	1
Tambaksari	1	1	1
Tandes	1	1	1
Tegalsari	1	3	1
Tenggilis Mejoyo	1	1	1
Wiyung	1	1	1
Wonocolo	1	1	1
Wonokromo	1	3	1

Lampiran 15. Hasil Pengelompokan FGK Kota Surabaya

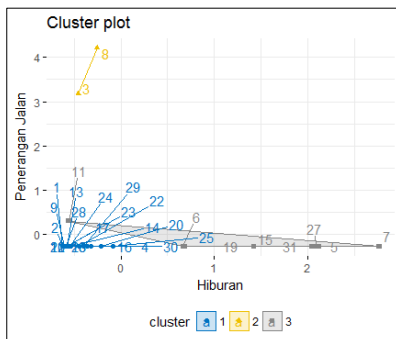
Kecamatan	2 Klaster	3 Klaster	4 Klaster
Asemrowo	2	3	2
Benowo	2	3	2
Bubutan	1	1	3
Bulak	2	3	2
Dukuh Pakis	2	3	1
Gayungan	2	3	1
Genteng	2	3	1
Gubeng	1	2	3
Gunung Anyar	2	3	2
Jambangan	2	3	2
Karang Pilang	1	1	4
Kenjeran	2	3	2
Krembangan	2	1	4
Lakarsantri	2	3	2
Mulyorejo	2	3	1
Pabean Cantian	2	2	2
Pakal	2	3	2
Rungkut	2	3	2
Sambikerep	2	3	1
Sawahan	2	3	2
Semampir	2	3	2
Simokerto	2	3	2
Sukolilo	2	3	2
Sukomanunggal	2	3	2
Tambaksari	2	3	2
Tandes	2	3	2
Tegalsari	2	3	1
Tenggilis Mejoyo	2	3	2
Wiyung	2	3	2
Wonocolo	2	3	2
Wonokromo	2	3	1

Lampiran 16. Plot Hasil Pengelompokan FKM Kota Surabaya

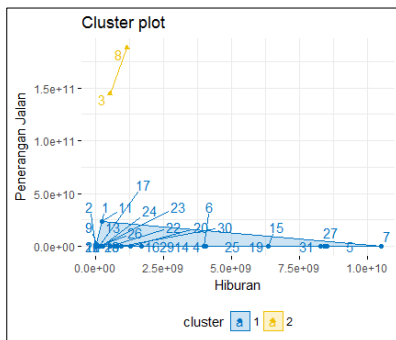
a. 2 Klaster



b. 3 Klaster

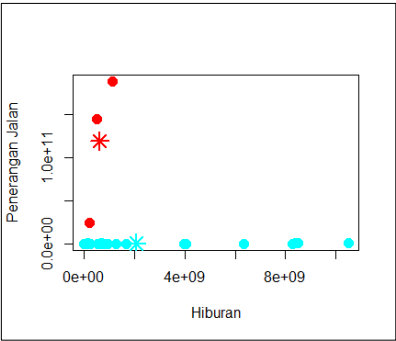


c. 4 Klaster

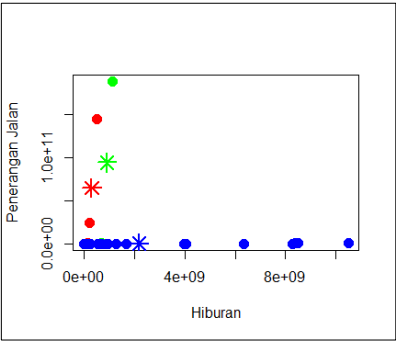


Lampiran 17. Plot Hasil Pengelompokan FGK Kota Surabaya

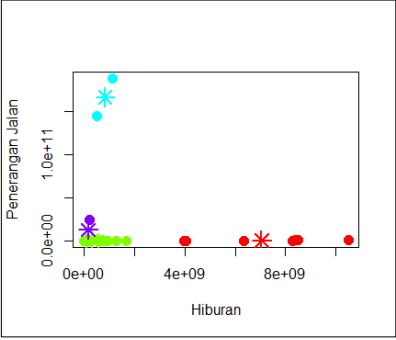
a. 2 Klaster



b. 3 Klaster



c. 4 Klaster



Lampiran 18. Syntax Uji Multivariat Normal

```
Macro
NormalMultivariate X.1-X.p qc dj2
MConstant i j n p Prop Tengah
MColumn x.1-x.p xj Kali d dj2 qc Prob
MMatrix MCova MCoval xjxbar
let n=count(x.1)
Covariance X.1-X.p MCova
invers MCova MCoval
do i=1:n
  do j=1:p
    let xj(j)=x.j(i)-mean(x.j)
  enddo
  copy xj xjxbar
  mult MCoval xjxbar Kali
  let d=Kali*xj
  let dj2(i)=sum(d)
enddo
sort dj2 dj2
do i=1:n
  let Prob(i)=1-(n-i+0.5)/n
enddo
INVCDF Prob qc;
  Chisquare p.
  corr qc dj2
  name qc 'qc'
  name dj2 'dj2'
endmacro
```

Lampiran 19. Syntax Pseudo F-Statistics dan Icd-rate

```
index.G1(df2016,df2016.fcm.naik.2k$clustering)
index.G1(df2016,df2016.fcm.naik.3k$clustering)
index.G1(df2016,df2016.fcm.naik.4k$clustering)
index.G1(df2016,df2016linear2.FGK$clus[,1])
index.G1(df2016,df2016linear3.FGK$clus[,1])
index.G1(df2016,df2016linear4.FGK$clus[,1])

df. = cbind(df2016,df2016.fcm.naik.2k$clustering)
df.
df.[df.$`df2016.fcm.naik.2k$clustering`==1,1:2]
ssw1 = (nrow(df.[df.$`df2016.fcm.naik.2k$clustering`==1,1:2])-
1)*sum(apply(df.[df.$`df2016.fcm.naik.2k$clustering`==1,1:2],2,var))
df.[df.$`df2016.fcm.naik.2k$clustering`==2,1:2]
ssw2 = (nrow(df.[df.$`df2016.fcm.naik.2k$clustering`==2,1:2])-
1)*sum(apply(df.[df.$`df2016.fcm.naik.2k$clustering`==2,1:2],2,var))
```

```

ssw = ssw1+ssw2
ssw
a = apply(df[, -3], 2, mean)
sst1 = sum(apply((df[, df2016.fcm.naik.2k$clustering`==1, 1:2] - a)^2, 2, sum))
sst2 = sum(apply((df[, df2016.fcm.naik.2k$clustering`==2, 1:2] - a)^2, 2, sum))
sst = sst1+sst2
sst
icdrate = 1-((sst-ssw)/sst)
icdrate

df.fgk=cbind(df2016, df2016linear4.FGK$clus[, 1])
df.fgk
df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==1, 1:2]
ssw1.fgk = (nrow(df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==1, 1:2])-
1)*sum(apply(df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==1, 1:2], 2, var))
df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==2, 1:2]
ssw2.fgk = (nrow(df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==2, 1:2])-
1)*sum(apply(df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==2, 1:2], 2, var))
df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==3, 1:2]
ssw3.fgk = (nrow(df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==3, 1:2])-
1)*sum(apply(df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==3, 1:2], 2, var))
df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==4, 1:2]
ssw4.fgk = (nrow(df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==4, 1:2])-
1)*sum(apply(df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==4, 1:2], 2, var))
ssw.fgk = ssw1.fgk+ssw2.fgk+ssw3.fgk+ssw4.fgk
ssw
a.fgk = apply(df.fgk[, -3], 2, mean)
sst1.fgk = sum(apply((df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==1, 1:2] -
a.fgk)^2, 2, sum))
sst1.fgk
sst2.fgk = sum(apply((df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==2, 1:2] -
a.fgk)^2, 2, sum))
sst2.fgk
sst3.fgk = sum(apply((df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==3, 1:2] -
a.fgk)^2, 2, sum))
sst3.fgk
sst4.fgk = sum(apply((df.fgk[df.fgk$`df2016linear4.FGK$clus[, 1]`==4, 1:2] -
a.fgk)^2, 2, sum))
sst4.fgk
sst.fgk = sst1.fgk+sst2.fgk+sst3.fgk+sst4.fgk
sst.fgk
icdrate.fgk = 1-((sst.fgk-ssw.fgk)/sst.fgk)
icdrate.fgk

```

Lampiran 20. Syntax FKM dan FGK Normal Multivariat

```
kova2i <- matrix(c(1,0.7,0.7,1),ncol=2,nrow=2)
miu <- c(90,90)
kotak1a <- mvrnorm(n = 750, miu, kova2i)
kotak1a
miu1 <- c(90,89)
kotak1b <- mvrnorm(n = 750, miu1, kova2i)
gedeya=rbind(kotak1a,kotak1b)
gedeya
mem2016.naik.2k = read.csv(file =
"D:/MF/MF1500/MF2016LinearNaik2K.csv",header=TRUE,sep = ";")
print(mem2016.naik.2k)
a10.2k = fanny(x=gedeya,k=2,memb.exp = 2,metric = "euclidean",iniMem.p =
as.matrix(mem2016.naik.2k))
print(a10.2k)
kesselya.FGK = FKM.gk(X = gedeya,m = 2,startU = mem2016.naik.2k, RS=50)
print(kesselya.FGK)
write.csv2(x=gedeya,file='D:/hasil/datanya/c10.csv')
write.csv2(x=a1.2k$clustering,file='D:/hasil/multivariat2k/fkm10.csv')
write.csv2(x=kesselya.FGK$clus,file='D:/hasil/multivariat2k/fgk10.csv')
```

Lampiran 21. Syntax FKM dan FGK Normal Univariat

```
seratuslimapuluh9a <- matrix(c(1,0.70,0.70,1),ncol=2,nrow=2)
miu9a <- c(90,90)
sq9a <- mvrnorm(n = 75, miu9, seratuslimapuluh9a)
sq9a
miu99 <- c(90,89)
sq99 <- mvrnorm(n = 75, miu99, seratuslimapuluh9a)
okedeh9a=rbind(sq9,sq99)
okedeh9a
write.csv2(x=okedeh9a,file='D:/hasil/datanya/e10.csv')
mem2016.naik.2k = read.csv(file =
"D:/MF/MF150/MF2016LinearNaik2K.csv",header=TRUE,sep = ";")
print(mem2016.naik.2k)
e9a.2k = fanny(x=okedeh9a,k=2,memb.exp = 2,metric = "euclidean",iniMem.p =
as.matrix(mem2016.naik.2k))
print(e9a.2k)
write.csv2(x=e9a.2k$clustering,file='D:/hasil/multivariat2k/fkm10.csv')
e9a.FGK = FKM.gk(X = okedeh9a,m = 2,startU = mem2016.naik.2k, RS=1)
print(e9a.FGK)
write.csv2(x=e9a.FGK$clus,file='D:/hasil/multivariat2k/fgk10.csv')
```

Lampiran 22. Syntax FKM dan FGK Surabaya

```
df2016 = read.csv(file="D://Data2016b.csv", header=TRUE, sep = ";")
df2016 = data.frame(df2016$Hiburan, df2016$PeneranganJalan)
colnames(df2016) = c("Hiburan", "Penerangan Jalan")
print(df2016)
mem2016.naik.2k = read.csv(file = "D:/MF2016LinearNaik2K.csv", header=TRUE, sep =
";")
print(mem2016.naik.2k)
df2016.fcm.naik.2k = fanny(x=df2016, k=2, memb.exp = 2, metric = "euclidean", iniMem.p
= as.matrix(mem2016.naik.2k))
print(df2016.fcm.naik.2k)
fviz_cluster(df2016.fcm.naik.2k, ellipse.type = "convex", repel = TRUE, stand = FALSE,
palette = "jco", ggtheme = theme_minimal(), legend = "bottom")
df2016.fcm.naik.2k$clustering
df2016linear2.FGK = FKM.gk(X = df2016, m = 2, startU = mem2016.naik.2k, RS=50)
print(df2016linear2.FGK)
plot(df2016linear2.FGK)
df2016linear2.FGK$clus
mem2016.naik.3k = read.csv(file = "D:/MF2016LinearNaik3K.csv", header=TRUE, sep =
";")
print(mem2016.naik.3k)
df2016.fcm.naik.3k = fanny(x=df2016, k=3, memb.exp = 2, metric = "euclidean", iniMem.p
= as.matrix(mem2016.naik.3k))
print(df2016.fcm.naik.3k)
fviz_cluster(df2016.fcm.naik.3k, ellipse.type = "convex", repel = TRUE, palette = "jco",
ggtheme = theme_minimal(), legend = "bottom")
df2016.fcm.naik.3k$clustering
df2016.fcm.naik.3k$crisp
df2016linear3.FGK = FKM.gk(X = df2016, m = 2, startU = mem2016.naik.3k, RS=50)
print(df2016linear3.FGK)
plot(df2016linear3.FGK)
df2016linear3.FGK$clus
mem2016.naik.4k = read.csv(file = "D:/MF2016LinearNaik4K.csv", header=TRUE, sep =
";")
print(mem2016.naik.4k)
df2016.fcm.naik.4k = fanny(x=df2016, k=4, memb.exp = 2, metric = "euclidean", iniMem.p
= as.matrix(mem2016.naik.4k))
print(df2016.fcm.naik.4k)
fviz_cluster(df2016.fcm.naik.4k, ellipse.type = "convex", repel = TRUE, palette = "jco",
ggtheme = theme_minimal(), legend = "bottom")
df2016.fcm.naik.4k$clustering
df2016linear4.FGK = FKM.gk(X = df2016, m = 2, startU = mem2016.naik.4k, RS=50)
print(df2016linear4.FGK)
plot(df2016linear4.FGK)
df2016linear4.FGK$clus
```

Lampiran 23. Surat Pernyataan Data

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMKSD ITS:

Nama : Yanuar Rafi Rahadian

NRP : 1313100091


menyatakan bahwa data yang digunakan dalam Tugas Akhir/ Thesis ini merupakan data sekunder yang diambil dari penelitian / buku/ Tugas Akhir/ Thesis/ publikasi lainnya yaitu:

Sumber : Badan Pengelolaan Keuangan dan Pajak Daerah Kota Surabaya

Keterangan : Data Realisasi Pajak Air Tanah, Pajak Hiburan, Pajak Penerangan Jalan, Pajak Hotel, Pajak Restoran, Pajak Parkir serta Pajak Bumi dan Bangunan

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui
Pembimbing Tugas Akhir


Raden Mohamad Atok, Ph.D.
NIP. 19710915 199702 1 001

Surabaya, 24 Januari 2017


Yanuar Rafi Rahadian
NRP. 1313100091

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Yanuar Rafi Rahadian, lahir di Surabaya 11 Januari 1995. Putra pertama dari dua bersaudara pasangan Rudy Setiawan dan Zamrud Mufida. Penulis menempuh pendidikan formal di SD Muhammadiyah 4 Pucang Surabaya; SMP Negeri 6 Surabaya; SMA Negeri 6 Surabaya. Penulis melanjutkan pendidikan Strata 1 Statistika FMKSD ITS dengan NRP 06211340000091. Selama perkuliahan, penulis aktif baik dalam organisasi dan komunitas. Selain menjadi *founder*

komunitas Bisnis Bareng Indonesia, organisasi yang pernah diikuti oleh penulis adalah menjadi anggota Departemen Penelitian dan Pengembangan Himpunan Mahasiswa Statistika ITS (HIMASTA-ITS). Dalam menambah jejaring relasi, penulis juga tergabung di komunitas Pengusaha Kampus Region Surabaya. Penulis pernah lolos tahap seleksi akhir *startup* di Start Surabaya dengan mempresentasikan game edukasi *Boyoboy* pada acara *d'preneur* yang diselenggarakan oleh detikcom. Saat ini penulis fokus menjadi pengajar lepas matematika tingkat SD dan dasar gitar pada lembaga bimbingan belajar *Speed* yang didirikan secara mandiri. Penulis melatih mengaplikasikan ilmu statistika dengan cara menjadi surveyor lepas MPM Honda dan Indomarco. Hal yang membahagiakan untuk penulis adalah ketika ilmu atau pengalaman yang dimiliki dapat menjadi kekuatan untuk menginspirasi dan menggerakkan orang lain untuk berkarya. Apabila pembaca ingin berdiskusi atau memberikan kritik dan saran dapat menghubungi penulis melalui email rafirahadian11a@gmail.com.